

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFG.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5839	99.3	1153	1 RWHUIB	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.8	1163	1 RWHUIC	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 S03308	lymphocyte fuction
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 A44142	integrin alpha-2 c
10	1054	17.9	1181	2 A33998	integrin alpha-1 c
11	1049	17.8	1180	2 A35854	integrin alpha-4 c
12	666	11.3	1039	2 A41131	integrin alpha-9 c
13	642	10.9	1035	2 S06046	integrin alpha cha
14	630	10.7	1035	2 I58409	alphaP integrin -
15	614.5	10.5	1041	2 T31437	cell surface glyco
16	579.5	9.9	1054	2 J07294	VLA-3 alpha subuni
17	572.5	9.7	1051	2 A35761	integrin alpha-5 c
18	567.5	9.7	1053	2 I55534	integrin alpha-V c
19	555.5	9.4	1053	2 S44250	integrin alpha-v c
20	551.5	9.4	1034	2 A36108	integrin alpha-6 c
21	539	9.2	1044	2 T10050	fibronectin recept
22	533.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	integrin alpha-5 c
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S16516	integrin alpha-8 c
29	497	8.5	1394	2 A29637	position-specific

ALIGNMENTS

RESULT 1

RWHUIB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1; integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C/Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; 152867

R/Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A>Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CR3)

B.

A/Reference number: A31108; MUID:88315033; PMID:2457584

A/Accession: A31108

A/Molecule type: mRNA

A/Residues: 1-1153 <COR>

A/Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307148

A/Note: part of this sequence was confirmed by protein sequencing

R/Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A>Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M

A/Reference number: A28915; MUID:88257215; PMID:2454931

A/Accession: A28915

A/Molecule type: mRNA

A/Residues: 1-499,501-965, 'P', 967-1153 <ARN>

A/Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594

A/Note: the authors translated the codon TAC for residue 1129 as Thr

A/Note: part of this sequence, including the amino end of the mature protein, was confir

R/Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A>Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A/Reference number: A41600; MUID:92073318; PMID:1683702

A/Accession: A41600

A/Molecule type: DNA

A/Residues: 1-9 <SHE>

A/Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215

R/Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A>Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi

A/Reference number: A94193; MUID:88190151; PMID:2833753

A/Accession: A30892

A/Molecule type: mRNA

A/Residues: 917-1042 <AR2>

R/Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A>Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A/Reference number: A32218; MUID:89098893; PMID:2563162

A/Accession: A32218

A/Molecule type: mRNA

A/Residues: 9-1153 <HIC>

A/Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

A;Note: Part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-1
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <PLE>
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA824821.1; PID:G263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:92144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:G180184; PIDN:AA51960.1; PID:G553219
A;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: Promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <Sig>
F;17-1153/Product: cell surface glycoprotein CD11b #status
F;17-1108/Domain: extracellular #status predicted <Ext>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSLGLSLAATSPPOLACGPTVHTQTSNTYVKGCLCFLFGSNLRQPPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPPOLACGPTVHTQTSNTYVKGCLCFLFGSNLRQPPQK 136

Qy 121 FPEALRGCPQEDSDAFIDGSGSIIPHDFRMKELVSTIMEOLKSKTLESLMOYSEEF 180
Db 137 FPEALRGCPQEDSDAFIDGSGSIIPHDFRMKELVSTIMEOLKSKTLESLMOYSEEF 196

Qy 191 RIHFTFKFQNNPNRSLKPIQTLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPNRSLKPIQTLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256

Qy 241 TDGEKFGDPLGYEDVPELDRGVTRYVLGFGDAPRSKSRQELNTVASKPPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSKSRQELNTIASKPPRDHVFOVN 316

Qy 301 NPEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSLVGSDVWAG 360
Db 317 NPEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLSLVGSDVWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPYQHIGLVAMPR 436

Qy 421 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCP 480
Db 437 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSVCP 496

Qy 481 PRGQARWQCDVAVLYGEQGPGRFGAALTVLGDVNGDKLTVDVAGAPEDNRGAVLYLF 540
Db 497 PRGQARWQCDVAVLYGEQGPGRFGAALTVLGDVNGDKLTVDVAGAPEDNRGAVLYLF 556

Qy 541 HFTSGSIGSPHSQRIAGSKLSPRLQYFQOSISGGQDLTMDGLVDLTGCAQGHVLLRSQ 600
Db 557 HFTSGSIGSPHSQRIAGSKLSPRLQYFQOSISGGQDLTMDGLVDLTGCAQGHVLLRSQ 616

Qy 601 PVLRVKAIMENPREVARNVFECNDQVKGKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVFECNDQVKGKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 736

Qy 721 SLVGTPLSAPGNLRPVLAEDAQRLFTALPFPEKNCNDNICODDLSITTFMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLRPVLAEDAQRLFTALPFPEKNCNDNICODDLSITTFMSLDCLVVG 796

Qy 781 GPREFNVTVTRNDGDSYRQTVTFPPDLDSYRKVSTLQNSQRSWRSLACESASSTEV 840
Db 797 GPREFNVTVTRNDGDSYRQTVTFPPDLDSYRKVSTLQNSQRSWRSLACESASSTEV 856

Qy 841 SGALKSTSCSINHIPIFENSEVTNITFDVDSKASLGKULLKANVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHIPIFENSEVTNITFDVDSKASLGKULLKANVTSNNMPTNKTEF 916

Qy 901 QLELPVKYAVYVMTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNTFASNTSRVMOHQYQVSNLQORSIPISLVFLVPV 976

Qy 961 RLQNTVWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020
Db 977 RLQNTVWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1036

Qy 1021 FFGIOEEFNATLKGNSLDFDWMYIKTSHNHLIIVSTABILFNDSVFTLLPGQAFVRSOTET 1080
Db 1037 FFGIOEEFNATLKGNSLDFDWMYIKTSHNHLIIVSTABILFNDSVFTLLPGQAFVRSOTET 1096

Qy 1081 KVEPFEPVNPPLIVGSSVGLLALLALITAAALKYKLGFFKROYKDMSEGPGPGEAPQ 1137
Db 1097 KVEPFEPVNPPLIVGSSVGLLALLALITAAALKYKLGFFKROYKDMSEGPGPGEAPQ 1153

RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; 159078
R;Pyteia, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PVT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G52982
A;Note: the authors translated the codon CAC for residue 569 as Gln

```
QY      781 GPREFNVTVTRNDGEDSVRTQTWTFPPDLISVRKYSTLQNORSQSRL-ACESASSTE   839  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       796 GFQDFNMSTVLRNDEGSDVGTVYYPSGLSYRKDSASNPLTKPKFWFKPAESSSSSE   855  
QY      840 VSGALKTSITCSINHPIFPENSEVTNNITFDVDVS-KASLGKKLLKANVTSENMPNKTE    899  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       856 GHGALSKSITWINIHFIIPANSEVTNTFDDVSHAFSGNKLKKAIVASENNMSRTHKT     915  
QY      900 FOLELPVKAVVMVVTTSHGVSTKYLNFTASENTSRVMQHQQVSNIGQRSLSFLVLVP    959  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       916 FOLELPVKIAYIMIVTDESIRYLNF-TASEMTSKIHQYQYFNNLGORSLPVS-VWFVIP    975  
QY      960 VLINQTVIWDRPQPVFSENLSNSTCHTKERLSPHSDFLAELRKAPVNVGSIAVCQRICDI   1019  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       976 VOINNVTWDHPQVIFSQNLSSACHTEOKSPHNSFRQLERTPVINGC-SVAACKRIQCCL   1035  
QY      1020 PFFGIQEENATLKGNISFDWIYIKTSHNHLLIVSTAETILFNDSVFTLLPGOGAFVRSOTE   1079  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       1036 PSFNTQEIFNYTLKNLSFDWKVIKTSHGHLILLVS-STELFNDSA-FALLPGOESYVRSKTE   1095  
QY      1080 TKVEPFEPVNPLELVGVGGMLLAILTAALYKLGFFKRYKDMMGEGFGPGAEPQ        1137  
Db      :|||:|||::|||::|||::|||::|||::|||::|||::|||:  
       1096 TKVEPYEVHNPEVLIVSGSIGGLVLLAILTAGLYKLGFFKRYKDMMNEA-PQAAPPQ    1153
```

RESULT 3

RWHUIC

N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004

C;Accession: A36584; A35543; S00864

J;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J; Biol. Chem. 265, 12750-12751, 1990

A;Reference number: A36584

A;Contents: erratum

A;Accession: A36584

A;Molecule type: DNA

A;Residues: 1-1163 <COR>

A;Cross-references: UNIPROT:P20702

A;Note: this revision to the sequence from reference A35543 includes the carboxyl end

F;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule

A;Reference number: A35543; MUID:90153906; PMID:2303426

A;Accession: A35543

A;Molecule type: DNA

A;Residues: 1-834 <CO2>

A;Note: this sequence has been revised in reference A36584

F;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J., 6, 4023-4028, 1987

A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A;Reference number: S00864; MUID:88166645; PMID:3327687

A;Accession: S00864

A;Molecule type: mRNA

A;Residues: 1-755,'L',757-1163 <CO3>

A;Cross-references: GB:M81695; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my

C;Genetics:

A;Gene: GDB:ITGX; CD11c

A;Cross-references: GDB:119758; OMIM:151510

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein Cd1lb; von Willebrand factor type A repeat homoc

K;Keywords: calcium, cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <Sig>

F;20-1163/Product: cell surface glycoprotein Cdl1c status predicted <MAT>

F;20-1107/Domain: extracellular #status predicted <EXT>

F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>

F;1108-1133/Domain: transmembrane #status predicted <TM>

F;1134-1163/Domain: intracellular #status predicted <INT>

F;1134-1163/Domai


```

QY 768 TFSFMSLDCLVGGP-----REENVTVTVRNDGSDSYRTQVTFPPFLDLSYRKVSTLQN 821
D 780 SSPARS-----GFLRLMSASLAVETWLSNGEDAYWRLDLPFRGLSFFKVEMLQ- 831
QY 822 QRSQRWRLACESASSTEVSGAL-KSTSCSINHIPIPEENSEVTFNITFDVDSKASLGNKL 880
D 832 --PHSRMPVSCBEL--TEGSSLLTKUKCNVSPGIFKAGEVSLQWNTLLNSMEDFV 887
QY 881 LLKANVTSEN-NMPTNKTFQLELPVKYAVYVVTSHGVSTKYLNFTASENTSRVMQHQ 939
D 888 ELNGTVHCENENSLQEDNSAATHIPVLYPNVILTKBOENSTLYISPTPKGPKTQOVQV 947
QY 940 YQVNSLQORSLPSLVLNPNVRLNQTVIMDRPQ-----VTFSENLS-----TCHTKE-RLP 990
D 948 YQV-----RIQPGAYDHNMPT-LEALGVVPRPHSEDLTITVWSQTDPLVTCHESEDKRP 1001
QY 991 SHSDFLAELRKAPVNVCSIAVCORIQCDIPFGIOEBFNATLKNLSFDMVYIKTSHNLL 1050
D 1002 SSE---AEQPCLPV-----QRCPIVP---RWELLQVTVGLVLSKEIKAS-STLS 1046
QY 1051 IVSTAELFNDSVFTLLFGQAFYRSQTEKVPPEVFPNPLPLIVGSGVGLLLALITA 1110
D 1047 LCSSLVSFNSKHFHLYGSKA-SEAQVLVKVDLHEKEMLHVTVLVSIGIGLVLLFLIFL 1105
QY 1111 ALYKLGFPKQYKMM-SEGPPGAEP 1136
D 1106 ALYKVGFPKRLKMEADGGVNGSP 1132

RESULT 6
A:Species: human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C:Accession: A53213
J:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J: Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit.
A:Reference number: A53213; PMID:94164962; PMID:8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: UNIPROT:P38570; GB:I25851; NID:G457244; PID:G457245
C:Genetics:
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.2%; Score 1128; DB 2; Length 1179;
Best Local Similarity 28.7%; Pred. No. 2.9e-70;
Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS-CEPI-RLQVP-----VEAVNMSLGLSAAITSPQLACGPTVHQ 95
D 65 GPLHRCSLVQDELCHPVEHVPKGRHGVTVVRSHGVLICI-----QVLVRRP--HS 117
QY 96 TCSENTYVKGCLFGLSGLNRQKQPO----- 119
D 118 LSESLT---GTCLLGLFDLRPQAQANFFDLNLLDPADVTDGDCYCNKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLDGSISIIHPDRFMKELVSTIMSQL-- 164
D 175 RQRALEKEEEDKEEEDDEEAEACTEAIILDGSIDPPDPFQAKDFISNMNRNFEY 234
QY 165 KSKTTLFSLMQYSEEFRIHTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFN 224
D 235 KCPECNFALVQYGVGTQTEFLRDSQDVMASTARVQNTQVSGVTKTASQHWLSDIFT 294
QY 225 ITNGARKNAKILFLITDGEKFGDPLGVEDVPELOREGVIRYVLGFGDAFRSEKSRQEL 284

```

```

D 295 SHGSRKSKVMMVLTDCGIFEDPLNLTITVNSPKMQGVVERPAIGVGEFFKASRTAREL 354
QY 285 NTVASPRPDHVPQANNFEALKTQVQNLREKIFAIEGTQTGTSSSSPHEHMSQGFSAAIT 344
D 355 NLITASPDDETHAFKVTNYMALDGLLSKLRIIISMEST--VGDALHYQLAQIGFSAQIL 411
QY 345 SNGP-LLSTVGSYDAGWGVFLY--TSKEKSTFINNTRVDSMDNA-----YLGYAAAILRN 398
D 412 DERQVLLGAVGDFWSGALLYDTRSGRFLNQTAATAAADAQAQYLYGYAVAVLHKT 471
QY 399 RVOSLVGAPRYOHIGLVAMFR-ONTGWNESNAV-KGTOIGAYFGFASLCSVDVDSNGST 456
D 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFPLVLEGEQMGVSFSGELCPVDIDMGST 529
QY 457 DLVLIGAPHYETRGQVSVCPRLPQRQARWQCDVLYGEGQGPWRFGAALTVLGDVN 516
D 530 DFLVAAAPFHVHGEGRVYVRLSE-QDGSFSLARILSGHPGFTNARFGFAMAAMDLS 588
QY 517 GDKLTDVAIGAP---GEEDNR--GAVLPHGTSGSISPSHSORIASGKLSPELOVFGQ 570
D 589 QDKLTDVAIGAPLEGFGADGASFGSVIYNG-HWDGLSASPSQIRIRASTVAPGLQVFGM 647
QY 571 SLSGQDLTMDGLVDLTVGQGHVLLRSOPVLKRVKAIMFENPREVARNVPCNDQVVK 630
D 648 SMAGGFDISGDGLADITVGLGQAVVFRSPVRLKVSMAFTPSALP-----IGF 697
QY 631 KEAGEVRVCLHVQKSTRDLREGIQSVTVYDLALDSGRPHSRVAFVNETKNSRRTQVL 690
D 698 NGVNVRLCFEI-SSVTTASSEGLREALNFTLDVGVKQRRRLQCSQVRSCLGCLREWS 756
QY 691 GLTQTCETLKLQLPN---CIEDPVSPVILNLSLVTPLSAFGLNLPVLAEDAQRLFT 746
D 757 SGSQLCEDL-LMPTEGELCEDCFSNASVSVYQL-QTPEGQTDHPQPIILDRTYEPFAI 814
QY 747 ALFPFKMCGNDNTICODDLSITFSMSLDCLVGSGPREENTVTVVRNDEGDSYRQVTF 806
D 815 FQLPYEKACKNKLFCVABLQLA-TTVSQOELVVGULTKELTLNINLTSGEDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKSTSCSINHPIFENSEVTNI 866
D 874 YPRNLQ-----LKEMQKPPSPNQCDDPQPV---ASVLIMNCRIGHPLV-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN-----MPTNKTEFQ---LELPVKAVYVTVTSHGV 919
D 924 VWQLEENAFNRTADITVTVNSNRRSLANETHTLQFRHGFAVLSKPSIMYVNTGQGL 983
QY 920 S--TKYLNFTASENTRVMQHOYQVSNLQORSLSPLVFLVPVRLNQTVIMDRPQVTFSE 977
D 984 SHHKEFLFVHGEN---LFGAEYQ-----LQICVTKLRLGLQVAAVKLTRLTQ 1028
QY 978 NLSSTCHTKERLPHSHSDFLAELRKAPVNVNCSIAVCORIQCDIPFGIOEBFNATLKNLS 1037
D 1029 ASTVCTWQSERACAYSS-VQHVVEHMSVSCVIA-----SDKENVTVAAEIS 1073
QY 1038 PDWYIKTSHNLLIVST-----ABILENDSVFTLLPGQAFVRSQTEKVPPEVFPNPL 1091
D 1074 WD-----HSELLKDVTELQILGETISFNKSIYEGNAENH--RTKITVTVFLKDEKXHS 1125
QY 1092 PLIVGSSVGLLALLITALYKLGFFKQYKDMSE 1128
D 1126 PIIIKSGVGLLIVLILVILFCGFFKRYQOOLNLE 1162

RESULT 7
A45226
Integrin alpha-1 chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45226
R:Briseswitz, R.; Epstein, M.R.; Marcantonio, E.E.
J: Biol. Chem. 268, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit

```


370 VGFSAEYSPQNNILMLGAVGAYDWSGTVVQKTPHGHILFISKQAFQEQILQDRNHSSVILGYS 429 Db
392 AATILNRVQSLVGLGAPRYOHIGLVAMFRONTGMBESNANV-----KGTQIGAYFGASL 445 Qy
430 VASISGTNSVHFVAGAPRANYTQIVLYSVN-----ENGNVTVIQSRGQDQIGSYFGSVL 484 Db
446 CSVDVDSNGSTDVLVLGAPHYBQTR--GGQVSCVCLPRQARARWQCDVILGEOQGPWG 503 Qy
485 CAVDVNDKTTIDVLLVGPAMYMDLKKEEGRVLTITIKG--ILNWH--QFLEPGNGLENA 541 Db
504 RFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLPHTSGSGISPSHSORLAGS--KL 561 Qy
542 RFGSATAALSDINMDGNDVIVGSPLENQNSGAVIYNGHEGM--IRLRYGQKILGSDRAF 600 Db
562 SPRLOYFGQSLSGQDITMDGLDVLTVGAQGHVLLRSQPVLRVKAIEMENPREVARNVF 621 Qy
601 SSHLOYFGRSLDGYDLNGDSITDVSAGFQVQVQLWSQSADVSVDASFTPKKI--TL 658 Db
622 ECNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVYDIALD---SGRPHSRAVEN 677 Qy
659 NKNAEI-----KLKLCF-----SAKFRPTQNNQVAIVNITIDEQFSRVSIRGLFK 707 Db
678 ETKNSTRTQTVLGLTOTCE--TLKLQLPNCIEDPVPSPVLRVLRNLSL--VGTPLSAFGL 733 Qy
708 ENNERCLOKTMVSAQACRSYIIHQEPS--DIISPLNLCMISLENPCT-----756 Db
734 RPVLAEADAQRLFTALFFPERKNCNDNICQDLSITP-----SFMSLDCLVWGGPREFNVTV 789 Qy
757 NPALAEYSETVKVPSIFPHKCGDGDGVCISDLVNVQLPATQOQPFIVSNQNRKLTFSV 816 Db
790 TVRNDGEDSYQTQVTFPPDLISYRKVSTLQNRQSRWRLACESASST-EVSGALKSTS 848 Qy
817 QLKNNKESAYNTEIIVDFSENLF-----ASWMPVDGTEVTQIASSQKSVT 864 Db
849 CSINHPIFENSEVTFTNFDVDSKASLGNKLLKANVTSENNMPRTNKTFFQLELPVKY 908 Qy
865 CNVGYPALKSKQCVTTFINFDNLQ--NLQONQASISFRALSESQENMADNSVNLKSLIY 923 Db
909 AVYMWVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPTISLVILV 958 Qy
924 DAEIHIT-RSTININFEVSLDGNVSSV-HSFE--DIGPKFISIKVTITGVSFVPSMA---976 Db
959 PVRLNQTWIDRPQVTFSEN---LSTCHTKE-----RLPSHSDPLAE- 998 Qy
977 -----SVLIHTPQYTKDKNPLMYLTGVHTDQAGDISCEAEINPLKIGOTSSSVFKSEN 1030 Db
999 LRKAPVWNCIAVCQRIQCDIPFGIOEEFNATLGNLSFDWYIKTSHNHLIVSTAET- 1057 Qy
1031 FRHIKELNRTASCNSIMCWLDRDLQVKGEYFLNVSTRINWNGTFAASTFQVLTAAAEID 1090 Db
1058 LFNDVSFTL-----LPGGAFVRSQTEPKVEFFE--VPNPLPLIVGSSVGGLLLLALITA 1110 Qy
1091 TYNPQIVIEENTVITP-----LTIMPKHEKVEVPTGVIVGSIAGILLLLALVA 1140 Db
1111 ALYKLGFFKRYQKDM 1125 Qy
1141 ILMKLGFFKRYEKM 1155 Db

RESULT 9
S44142

VLA-2 protein homolog - mouse
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S44142
R;Edelman, J.M.; Chan, B.M.; Uniaval, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A;Reference number: S44142
A;Accession: S44142
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:9473098; PIDN:CAA82877.1; PID:9473098
P;169-344/Domain: von Willebrand factor type A repeat homology <WMA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;
Best Local Similarity 27.8%; Pred. No. 27e-65;
Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;
QY 1 ENLDTENAMTFQ--ENARGFGQSVVQL---QGSRVVVGAPQEIIVAANQSGSLYOC--DYST 54
DB 27 YNVGLFGAKIFSGPSSEFGYVQQLTNPQGNLLVGSWSPGPPENRGMGVYKPCVDLPT 86
QY 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSENTYVKGCL 107
DB 87 ATCEKLNQNSASISNVTEIKTNMSLGLITRNPGTGGTFCGLPMAHQAQGNQYATGIC 146
QY 108 FLGSLNRLOQPO---KPEALRGCPQEDSDIAFLIDGSGSIIIPHDFFRMKELVSTIMEQL 164
DB 147 ----SDVSPDFQFLTSPSPAVQACPSL--VDVVVVCDESNSIYP--WEAVKNFLVKFTGL 199
QY 165 K--KSKTLFSLIMQYSEEFRIHFTFKFQNNPNRSLIKPTQLLQ--RTHATGLRKVVRE 221
DB 200 DIGPKTKQVALIQVANEPRIIIFNLNDFETKEDMVQATSETRQHGGLTNTFRAIEFARDY 259
QY 222 LFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDRGVTRY---VLGE--GDAFR 276
DB 260 AYSQTSGRPGATKVMVWVTDGESH--DGSKLKTVIQCNDDDEILIRFAGVIAVLGNRNALD 318
QY 277 SEKSRQELNTVASPPRDHVFQANNFALKTVQNLREKIFAIEGTQTGSSSSPHEMSQ 336
DB 319 TKNLKEIKALASTPTTERIFENVADAALKEKAGTLGQIIFSIETGVQ--GDNFQMEWAQ 377
QY 337 EGFA--AITSNGPLLSVGSYDWAGGVFLVTSKSTFTINMT--RVSDMN--DAYLGYA 391
DB 378 VGFSAIYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIFPKQAFQVQLQDRNHSSFLGYS 437
QY 392 AATILNRVQSLVGLGAPRYOHIGLVAMFRONTGMBESNANV---KGTQIGAYFGASLCS 447
DB 438 VAAISTEDGVHFVAGAPRANYTQIVLYSVNK---QGNVTVIQSHRGQIGSYFGSVLCS 494
QY 448 VDVSNGSTDLVLGAPHYBQTR--GGQVSCVCLPRQARARWQCDVILGEOQGPWGPRF 505
DB 495 VDVKDQTTIDVLLVGPAMYMDLKKEEGKVLFTITKILNQHQ---FLEGPEGTGNARF 551
QY 506 GAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLPHTSGSGISPSHSORLAGSKLSPR- 564
DB 552 GSAIAALSDINMDGNDVIVGSPVENNSGAVIYNGHQT--IRTKYSQKILGSGAFER 610
QY 565 -LOYFGQSLSGQDITMDGLDVLTVGAQGHVLLRSQPVLRVKAIEMENPREVARNVFC 623
DB 611 HLQFFGRSLDGYDLNGDSITDVSIGALGQVQLWSQSADVAIEALFTP-----660
QY 624 NDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVYDIALD---SGRPHSRAVENET 679
DB 661 -DKITLLNKDAKITLCLCFRAEFRAQONQV--ALLFNMTLDADGSHSRKVSRGVREN 717
QY 680 KNSTRRTQVGLGTQTCET--IKLQLPNCIEDPVPSPVLRVLRNLSLVTGTFSAFGLNRPVL 737
DB 718 SERFLQKNVWVNEVQKSEHHISIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768
QY 738 AEDAQRLFTALFPFEKNCNDNICQDLSI-----TFGFMSLDCLVWGGPREFNVMT 788
DB 769 EAYSETVKVPSIFPYKEGSDGICISDLILDVQQLPAIQTSF-----IVSNQNKLTFS 823
QY 789 VTVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASST-EVSGALKST 847
DB 824 VILKNKGESAYNTVILAESENLF-----AGSPMPVDGTEVTCEVGSQKSV 871
QY 848 SC SINHIPIFENSEVTFTNFDVDSKASLGNKLLKANVTSENNMPRTNKTTE--FQLELP 905
DB 872 TCDVGYPALKSEQCVTTFINFDNLQ--NLQONQAINFQAFSESQ--ETNKADSVSLTIP 928
QY 906 VKYAVYMWVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPTISLV 955

```

Db 929 LLYDAELHLLT-RSTNINFEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSWA 984
QY 956 FLV-----PVLNQTVIWRDPQVTF-SENLS 980
Db 985 LVTHIPIQYTKENPLLYLTGLTQDQAGDISCTABINPLKLPHTA-----PVSFKENFR 1040
QY 981 STCHTKERLPSHDSFLAELKAPVNVCSIAVCQRIQCDDIPFFGIOEENFATLKGNSPDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMKAEYFINVTVRWRT 1080
QY 1041 YIKTSHNHLIVTAEILFNDSVFLPLPGQAFVRSQTETKVEPPEVNPPLIVGSSVG 1100
Db 1081 FAASTFQTVQTAAREIDTHNPQFVIBENAVTIPLMIMKTEKAEVPT--GVIIGSIIA 1138
QY 1101 GLLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAMTAGLWKLGFQYKYM 1163

RESULT 10
A33998
Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:PR07301; GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catimel, B.; Parmentier, S.; Leung, L.H.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIb an
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutcher, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16,'V',18-21 <ZUT>
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 17.9%; Score 1054; DB 2; Length 1181;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

QY 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQVEIVANORGSLYQC--DYST 54
Db 30 YNVGLPEAKIFSGPSEQFGYAVQVQFINPKGNWLLVGSFGFPENRMDGVKCPVDLST 89
QY 55 GSCBPRLQ-----VPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCSNTYVKGLC 107

```

```

Db 90 ATCKLNLQTSSTIPNVTEMKTNMSLGLILTRNMGTCGLTCLPGLWAQCCGQNYTTGVC 149
QY 108 FLFGSNLRQQQKPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLK-- 165
Db 150 SDISPDF-QLSASPSAPQPCSL-IDVVVCDSDNSIYPWD--AVKNFLEKFFVQGLDIG 205
QY 166 KSKTLFSLIMOYSEEFRIHFTFKFQNNPNRSLIKPITQLLG-RTHATGRLKVVURELEN 224
Db 206 PTKTQVGLIQVANNPRVVFNLTVYKTEEMIVATSQTSQYCGDLTNTFGAIQYARKAYS 265
QY 225 ITNGARKNAPKLELLTDGKFGDPLGYEDVIPEDREGVIRY---VLGF--GDAFRSEK 279
Db 266 AASGRRSATKVVVVVTDGESH-DGSMKKAVIDQCHDNILRFGIAVLGLNRLALDTKN 324
QY 280 SRQELNTVASKPRDRHVFOANNFEALQVQNLREKIFAIEGTQTGSSSSPEHEMSQGF 339
Db 325 LIKEIKAIASIPTRYFFNVSDAALLEKAGTIGEIFSEGTVQG-GDNFQMEMSQVGF 383
QY 340 SAAITSNRP--LLSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGAAAI 394
Db 384 SADYSSQNDILMLGAVGAPCWSGTIVQKTSHGHLIFPKQAFDQLQDRNHSVLYGVAA 443
QY 395 ILNRVQSLVILGAPRVOHIGLVAMFRONTOMWESNANV-----KGTQIGAYGASICSV 448
Db 444 ISTGESTHFVAGAPRANYTQIIVLYSVN-----ENGNITVIAHRGDIQSGYFGLVCSV 498
QY 449 DVDSNGSTDLVLICAPHYEQTR--GGOVSVCPLPRQARWQCDVLYGEOQOPGRFG 506
Db 499 DVDKDTITDVLIVGAPWYMSDLKEEGRVLYFTIKKILGQHQ---FLEGEPIENTRF 555
QY 507 AALTVLGVNGDKLITVAIGAPGEEDNRGAVILFHTGSGSGISPHSORSOTAGS--KLSPR 564
Db 556 SAAIALSDIMDGFNDVIVGSPLENQSGAVIYNGHQT-ITKYSQKILGSDGAFRSH 614
QY 565 LQVFGOSLGGQDLTMDGLVDLTVAQGHVLLRSQPVLRVKAIMFNPREVARNVFECN 624
Db 615 LQVFGSLDGYGDGLNGDSITDVSIGAFGVQVQLWSQSIADVAIEAFTPEKI--TLVNKN 672
QY 625 DQVVKGEAGEVRVCLHVQKSTRDLRREGIQSIVVTDIALD-----SGRPHSEAVNETK 690
Db 673 AQII-----LKLCE---SAKFTPKONNQAIVINITLDADGFSSTVSRGLFKENN 721
QY 681 NSTRQTVLGLTQC--ETLKLQPLNCIEDVSPVILRNFSLVGTPLSAFGLRPLVIA 738
Db 722 ERLCKNWNVNAQSCPEHIIYIQPS--DVVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALPEKKNCGNDICQDDLSITF---SFMSLDCLVGGPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFHKDCEDGGLCISDLVDLRQIPAAQEPPIVSNQNKRLTFSVTLKN 832
QY 795 GEDSYRTQVTFPPLDLSVRKVSTLQNRQSRQSWRLACESAST-EVSCALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSVACDVG 880
QY 854 PIPENSEVTNITFDVDSKASLGNKLLKARVTSNNMPTNKTEFOLELPVKVAVMV 913
Db 881 PALKREQQVTFITFDNLQ-NLQNAQSLSCALSESQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSCHVSTKYLNFTASENTSRVMOHQVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLTRSTNINFYEISDGNVPSIVHSFEDVGPKFIFSLKVTVTGSPVSMATVIIHPQ 995
QY 959 -----PURLNQTIVWRDPQVTF-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLUKIQT-----SSVSFKSENFR---HTKE- 1047
QY 989 LPSHSDFLAELKAPVNVCSIAVCQRIQCDDIPFFGIOEENFATLKGNSLFDWIKTSHH 1048
Db 1048 -----LNCRTASCSNVTCWLKDVHMKGEYFVNTVTRNCTFASSTFQT 1091
QY 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRSQTETKVEPPEVNP-----LP--LIVGSSVG 1100

```

Db 1092 VOLTAAAEINTVPEIYVI-----EDNTVPIPLIMKPEDEKAEVPTGVIIISITA 1141

Qy 1101 GILLALLALITAAALYKLGFFKRYQKDM 1125

Db 1142 GILLALLALVAILWKLGFKKRYEKM 1166

RESULT 11

A35854

integrin alpha-1 chain precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004

C;Accession: A35854; S11243

R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990

A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A

A;Reference number: A35854; MUID:90338125; PMID:2380249

A;Accession: A35854

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1180 <IG>

A;Cross-references: UNIPROT:P18614; GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494

C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

F;1/70-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 17.8%; Score 1049; DB 2; Length 1180;

Best Local Similarity 27.3%; Pred. No. 9.9e-65;

Matches 344; Conservative 200; Mismatches 476; Indels 240; Gaps 48;

Qy 1 FNLDEAMTFQENARG-EGOSVVLQ---QGSRVVGAQPIVAANQSGSLYQCDYSTGS 56

Db 29 FNDVKNSMSGSPGVDMFGVTVQOYENEEKWVLGSLPLVGQPKARTGDVYKCPVGRER 88

Qy 57 CEP-IRLQVPVEA-----VNMSLGLSLAATTPSPQLLACQTPVHQTCSNTYVKG 106

Db 89 AMPCVKLLDFVNTSPNTEIKENWTFSTL-VINPENGFLACGLPVAAYRGCHLHYTTGI 147

Qy 107 CFLFGSNLRQOPKQPEALRGCPQSDIAFLIDSGSIIIPDPRMKELVSTIMEQLK 166

Db 148 CSDVSPTQVNVSEFAP---VQECSTQ-LDIVIVLDGNSIYP-----WESVIAFLNDLKR 199

Qy 167 -----SKTFLSLMOYSEFRHFTEKFEONPNRSLIKPITQLLG-RTHATGLRKKVR 220

Db 200 MDIGPKQTVGLVQGVNTEHFNLNKYSSTEEVLVAANKRGQGLQTMALGIDTARK 259

Qy 221 ELFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDRGVIRY---VLFGDA--F 275

Db 260 EAFTEARGARRGVKKVMIVITDGEH-DNYRLKQVIQDCEDENIQRFSTAILGHYNRGL 318

Qy 276 RSEKSRQELNTVASKPPRDHVPQANFEALKTQVQLREKIFAIEGTQTGSSSFEHMS 335

Db 319 STEKVEETKSTASEFTKHFNVSDDELALVIVKALGERIFALEBATDQSAASEMENS 378

Qy 336 QEGFSAATNSGPLASTVGSYDAGGVFLYTSKEXSTFINMT--RVDSDMD--AYLGY 390

Db 379 QTGFSAHVSQDMWMLGAVGAYDNGTVMQKANQWVPHNTTFQTEPAKNEPLASYLG 438

Qy 391 AAAAILNRVQSLVILGAPYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSDV 450

Db 439 TVNSATIPGDLVYIAGQPRYNTHTQVQVYKMDGNGINILQTLGGEQIGSYFGSVLTIDI 498

Qy 451 DNGSTDLVLIGAPHY-----YEOTR-GGQVSVCPPLRGORARWQCDAVL 494

Db 499 DKDSYTDLLVGLAPYMTGMTKEEKGKVVYVAVNQTTFEYQMSLEPRTQTCSSLKNSCT 558

Qy 495 YGFQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILFHTGSGSISPSHS 553

Db 559 KENKNEPCGARFGTATAANKDLNVDGNDVWIGAPLEDDHAGAVIYHG-SGKTIREAVA 617

Qy 554 QRIAGSKSLPRLQFYQSGGQDLTWDGLVDLTVGAAQGHVLLRSQPLRVKAIKFEFP 613

Db 618 QRIPSGGDGKTLKFFGQSIHGEDMDLNGDLTDTVIGLGGAAALFWARDVAVVKVMTNFP 677

Qy 614 REVARNVFECNDVVVKGEAG--EVRVCLHVQ-KSTRDLREGQTSQVVVYDLALDGRP 670

Db 678 NKVNIQKNCR---VEGKETVCINATMCPHVKLSKESDIYADLQ---YRVTLDSLQ 730

Qy 671 HSRVAFNET-----KNSTRQOTVGLGTQTCETLKLQLPNCI-----EDPVSPIV 715

Db 731 ISRSFPFGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVR 773

Qy 716 LRLNFSLVGTPLSAFCNLRPVLAEADAQRLFTALFFPEKNCNDNI CODLSITFSFMSLD 775

Db 774 VTLDFELT-DPENG-----PVLDDALPNSVHEHIPAKDCGNKERCISDLTINVSTTEKS 827

Qy 776 CLVVGSGPRE-FNVTVTVNRNDEGSYRQTQTFPPDLDSYRKVSTLQNRQSRQSWRLACES 834

Db 828 LLIVKSQHDKFNVLTVKNGDSAYNTRTVVQHSPLIFSGIBEIOKD-----SCES 879

Qy 835 ASSTEVSGALKSTSCSINHIPIFENSEVTFNITFDVDSKASLGNKLL-LKANVTSENMP 893

Db 880 N-----ONITCRVGYFPFLRAGETVTFKILFQFNWYSHLSENAIHLSATSDSEEPLE 930

Qy 894 RTNKTEFQLELPVKYAV---YMWVTSHGVST-----KYLNFTASENTSRVMQHQYQV 942

Db 931 SLMDNEVNISIPVKYEVGLQFYSSAHEHISVAANETIPEFINST--EDIGNEINVEYTI 988

Qy 943 SNLQORSPL--ISLVP-----LVPVRLNQTVMD-----RP----- 971

Db 989 RKGHGFMPPELQLSISFPNLTADGYPVLYPIG-----WSSSDNVNCRPSLEDPFGINS 1042

Qy 972 --QVTFPS-----ENLSSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORICDI 1019

Db 1043 GKMTITSKSEVLKKGTTIQDCSSTC-----GVATITCSLLPSDLSQVNV 1085

Qy 1020 -----PFFGIQBERF---NATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQ 1070

Db 1086 SLLLWKTFP-IRAHFSSNLTLRGELK-----SENSSLTSSN----- 1123

Qy 1071 GAFVRSQTEKVEFEFFVNPPLPL--IVGSSVGGLLALLALITAAALYKLGFFKRYQKDMSE 1128

Db 1124 --RKREIAIQISKDGLPGRVPLWVILLSAFAGLLMLLALLALWKGIFRKLKKMEK 1180

RESULT 12

A41131

lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N;Alternate names: integrin alpha-4

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C;Accession: A41131; S16742

R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L. J. Cell Biol. 115, 1149-1158, 1991

A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte A;Reference number: A41131; MUID:92064645; PMID:1840602

A;Accession: A41131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1039 <NEU>

A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51484

C;Superfamily: integrin alpha-4 chain

C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 666; DB 2; Length 1039;

Best Local Similarity 22.8%; Pred. No. 5.3e-38;

Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48;

Qy 1 ENLDTENAMTFQ-ENARGFGQSVV-QLOGSR--VVVGAPOEIVAA--QREGSLYQCDY 52

Db 41 YNLDPENALLYQPSGTLFGYVVLHSHGSKRWLIIVGAPTASWLSNASVNVNPGAIYRCGI 100

Qy 53 STG---SCFIRLQVP-----VEAVNMSLGLSLAATTPSP-PQLLACG---PTVHQ 95

Db 101 RKNPNQTCQLQSGSPSGPCGKTCLEERDNOMLGVTLFRQPGENGSIIVTCGRHWKNIFY 160

Qy 96 TCSNTYVVKGLCLFLFNSLRQOPKQPEALRGCPQSDIAFLIDSGSIIIPDPRMKE 155

[illegible]

Db

984 TIIITISLLGLIVLLILSCVMWKAGFFKQYKSILQE 1022

RESULT 13
S06046

integrin alpha-4 chain precursor - human
N;Alternate names: Cd49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C;Species: Homo sapiens (man)
C>Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S06046; A39355; D28018
R;Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989

A>Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A;Reference number: S06046; MUID:89356603; PMID:2788572
A;Accession: S06046
A:Molecule type: mRNA
A;Residues: 1-1038 <TA>
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A>Title: Characterization of the alpha-4 integrin gene promoter.

A;Reference number: A39355; MUID:91239513; PMID:2034655
A;Accession: A39355
A:Molecule type: DNA
A;Residues: 1-93 <ROS>
A;Cross-references: GB:M62841
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A>Title: The very late antigen family of heterodimers is part of a superfamily of molecu
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: D28018
A:Molecule type: protein
A;Residues: 40-50,'E','52-'S3 <TA2>
C;Genetics:
A;Gene: GDB:ITGA4; CD49D
A;Cross-references: GDB:I28032; OMIM:192975
A;Map position: 2q31-2q32
C;Superfamily: Integrin alpha-4 chain
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmem
F;1-39/Domain: signal sequence #status predicted <SIG>
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.9%; Score 642; DB 2; Length 1038;
Best Local Similarity 23.6%; Pred.No.2.5e-36;
Matches 293; Conservative 170; Mismatches 408; Indels 368; Gaps 51;

QY 1 FNLDTENAMTFO--ENARGFGSQVV-QIQGSR---VVGAPQEIIVAAN-----QRGSLYQCDDY 52
 :|::||::: : : ||||| |::| |
Db 40 YNVDTESALLYQGPHNTLFQYSVVLHSHGANRWLLVGAPTANWLANVASVINPGAIYRCRI 99
 :|::||::: : : ||||| |::| |

QY 53 STG----SCBPIRLQVFVEAVNMNSLG:SLAATTSPPQLLAGCPHVHOTCSENTYVKGLCFI 109
 :|::||: : |::| |::| |::| :
Db 100 GKNPQTCEQLQIGSP-----NGEP---CG-----KTCLERDNQWL--- 133
 :|::||: : |::| |::| |::| :

QY 110 FGSNLRQQPQKPFEALRGCPQEDSDIAFLDGSGSIIPHDFRMKLVISTIMEOLKKSKT 169
 :|::||: : |::| |::| |::| :
Db 134 -GVTLRSOPGE-----NGSVTCGHR-----WKN 156
 :|::||: : |::| |::| |::| :

QY 170 LFSLMQYSEEFRIHFTTFKEFNPNPRSLIKPTQLLGRTHTATGLRKRVRELFININGA 229
 :|::||: : |::| |::| |::| :
Db 157 IFYI----- 160

QY 230 RKNAPKILLTDGKFQDPGLGYDIVPELDREGVIR-----YYLVGFGDARPERSEKRQ 282
 :|::||: : |::| |::| |::| :
Db 161 -KNENK-----LPTGCYCVP-----PDRLTELSXRIAPCYQDYVKFKGENPAS----- 203
 :|::||: : |::| |::| |::| :

QY 283 ELNTVASKPPRDHDVFOANNFEALKVTQNQLREKIFAIEGTQTGSSSFHEMSQEGFSAA 342
 :|::||: : |::| |::| |::| :
Db 204 -----COAGISSP 211

QY 343 ITSNGPILLSVGSVDWAGGVFLY---TSKEKSTFINMTRVDSMDNDAYLGYA-AAAILR 397
 :|::||: : |::| |::| |::| :
Db 212 YTDLIVMGAFGSSVWTGSLFVNITTNYKAFLDKQNKQVKF---GSYLGSYGAGHFRRS 268
 :|::||: : |::| |::| |::| :

R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993

A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partner
A;Reference number: A49459; MUID:94064789; PMID:8245132

A;Accession: A49459

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 10-1035 <PAL>

A;Cross-references: GB:L24158

C;Superfamily: Integrin alpha-4 chain

C;Keywords: glycoprotein; metal binding; transmembrane protein

F;1-27/Domain: signal sequence #status predicted <SIG>

398 NRVQSLVGLAPRQHIHGLVAFRQNTGMWESNANV---XGTOIGAYFGASLCSVDVDSN 453

269 QHTTEVVGAPQHEQIGKAVIF---SIDEKELNHLHMKKKLGGYFGASVCAVDLND 324

454 GSTDLVLIGAPHYEQTRGGQVSCVPLPRQORARWQC-DAVLYGEQOGPWGRFGAALTVL 512

325 GFSDL-LVGAPMQSTIREGRVFY-INSGGAVMMAMETNLVSGDKYA-AREGESLVNL 381

513 GDVNGKLTDTVAIGAPEEDNRGAVLFGHTSGSGISPSHSQRIAGSKLPRIQYFQOQL 572

382 GDINDGDFEVALGAPEDDLQGAIIYNGRA-DGISSTFSQRLEGIQSKLSMFQSI 440

573 SGGQDLTMDGLVDLTGGA--QGHVLLRSQPLVLRKAIMFNPREVARNVFECDNQVVK 630

441 SGQIDANNGVVDVAVGAFRSDSAVLLRTPVVIDASLS-HPESVNRKTFDC----- 492

631 KEAGEVRVCLHVQKSTRDRLRREGQIQSVVYDLDLDSGR-----PHSRVFNKSTRR 685

493 VENGWPSVCIDLTLFCYKGEVPGYIVLFYNNMSLDVNRKABSPRFFYFSGNGTSVITG 552

686 QTVGLGTQCTETLKLQPNCTEDPSPVILRLNFSLVGTPLS-----AFGNLRPVLAE 740

553 SIQVSSREANRTHQAQMRKQVDRILTIPIQIEAAYHLGPHVISKRSTEEFPPLQPLQOK 612

741 AQR-LFTALPFFKNCNGNDNTCODLSIT--PSFM---SLDCLVGGPREFNVTVVRN 793

613 KEDIMKKTINFARFAHEN-CSADLQVSAKIGFLPKPHENKTYLAIVGSMKTLMLNVSFN 671

794 DGEOSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASTEVSGALKSTSCSINH 853

672 AGDDAYETTLHVKLPGVLYFKILELEBK-----QINCE---VTDSNGVQ-LDCSLG 721

854 PIFPENSVTNIPFDVDSKASLGNKLLKANVTISEN--NMPRTNKTFFQLELPVKYAVY 911

722 IYVDHLSDIDSLFDVSSLSRAEEDLSITVHATCENEEDNKLKSHRVTVVAIPLKVEVK 781

912 MVVTSHGYSKYLNFTASNTSRV-----MOHQYQVSNLQGRSLP-ISLVLFPVRLNQ 965

782 LTVHGFSVNPSTFVYSGSDNENPETCWEKMKMLTFHVINTGSMAPNVSVIENVPSFS-- 839

966 VIWDRPQVTFENL-----SSTCHTKERLPSHSDFLAELRKAAPVNCVSIACVQRIQCDI 1019

840 -----PQTDKLFNILDVTTGECHEFNYQ-----RVCALEQ--- 871

1020 PFGIQEENATLKNLGDWDYIKTSHNHLIVSTAE-----ILFN-----DSVPT 1065

872 -----QKSNWQTLKGVRE---LSKTDKELLYCIKADPHCLNFCNFGWMSGKEASVHI 923

1066 LLPGGQAFVRSQTEKVPPEV-----PNP----- 1090

924 QLEGRPS-ILEMDETSALKFEIRATGFPPEPNEPVIENLKNDEVAHVLLLEGLHHQRPKRYF 982

1091 -IPLIVGSSVGLLILALITAAIYKLGFKPKROYKOWMSE 1128

983 TIVIISSSLGLGLVILLISYVNWKAGFPKRYKSLQE 1021

RESULT 14

I58409

Integrin alpha-9 chain precursor - human

C;Species: Homo sapiens (man)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004

C;Accession: I58409; A49459

R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah

Oncogene 9, 611-619, 1994

A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small

A;Reference number: I58409; MUID:94119603; PMID:8290272

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1035 <RES>

A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:G464180; PIDN:BAA04984.1; PID:G533327

Query Match 10.7%; Score 630; DB 2; Length 1035;

Best Local Similarity 26.5%; Pred. No. 1.8e-35;

Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

331 EHEMSQGFSAATISNGPPLLSTVSGYDWAGGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389

190 EHSGCOAGIAGFTTEELVVMGAPGFYWAGTIKVLNLT-DNTYLLKLNDEVMNRRYYTLG 248

390 YA-AAIILNRVQSLLVGLAPRYCHICLVAMER--QNTGMWESNANVKGTOIGAYFGASL 445

249 YAVTAGHFSHPSTIDVVGAPQDKGIGKVIFRADRRSGTLIKIFQASGKMGWVFGSSL 308

446 CSVDVDSNGSTDLVLIGAPHYEQTRGGQVSCVPLPRQORARWQCDVLYGEQGPWGRF 505

309 CAVDLNGGLSDL-LVGAPMFESEIRDEQVTVY-INRNGALEE-QLALTGD-GAYNAHF 364

506 GAALTIVLGDVNGDKLTDVAIGAPEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRL 565

365 GESTASLDLDDNDGFPDVAIGAPKEDDPAGAVIYHGDAG-GIVPQYSMKUSGGKINPVL 423

566 QYFGOSLGGQDLTMDGLVDLTGGA--QGHVLLRSQPLVLRKAIMFNPREVARNVFE 623

424 RMFGOSGGIDMDGNGYDPDVIAGFMSDSVLLRARVPIVD-VSIFLPGSINITAPQC 482

624 NDQVVKGEGEVRV-CLHVQ-----KSTRDLRREGQIOSVVTYDIALDSGRPHSRV 675

483 HD-----CQQPVNCLNVTTCSFHGKHVPEEI---GLNVLMADVAKKEGQMPRVY 531

676 F---NETKNSRTRQVIGLTQCTETLKLQPNCTEDPVSPIVLELNFSL---VGTPLS 728

532 FVLLGETMGQVTEKQLTYMBETCRHYVAHVKKRVQDVISPIVFEAAVLSLSHVHTGEEER 591

729 AFGNLRPVL-----AEDAQRULTALPFEKNCNGNDNICODDLSITFSFM--SLD----CL 777

592 ELFPPLTPVLRWKKGQKIAQKQTV---FERNCRSED-CAADLQLOGKLLSSMDEKTLVL 647

778 VVGGPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASS 837

648 ALGAVKNISLNIISNLGDDAYDANVSFNVSRLEFP-----INMWQKEEMGISCELDES 701

838 TEVSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNK 897

702 DFL-----KCSVGFPFMRSKSYEFVIFDTSHLS--GEEVLSFVITAQSG--NTER 750

898 TE-----FQLELPVKYAVYVWVTS-----HGVSTKYLNFTASENTS---RVMQHOY 940

751 SESLHDNTLVLMVPLMHEVDTSITGMTSPTSIFYGESVDAAFIQLDDLECHQFQPINITL 810

941 QVSNLQGRSLPISLVLEL-VPVRLN-----QTVIWDPRQVTFSENLSSTCHTKER 988

811 QVYNTGPSTLPSSVSISFENRLSSGGAEMFHVQEMVGVQEKGNCSFQKNPTPCIIPQEQ 870

989 LPSHSDFLAELRKA-----PVVNCVSIACVQRIQCDIPPPGIGOEENATLKNLSFD 1039

871 ENIFHTIFAFFTKTSGRKVLDCXPKGISCULTAHCN-----FSALAKEESRTI-----D 917

1040 WYIKTSHNHLIVSTAEIILFNDSVFTLLPQGGAFVRSQTEKVP----- 1084

918 IY-----MLNLT-EILKKDSSVIO-----FMSRAKVKVDPALRVVVEIAHGNPEEV 962

Search completed: November 9, 2004, 12:19:22
Job time : 34.25 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481b-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5823.5	99.1	1152	1 ITAM HUMAN	P11215 homo sapien
2	4447	75.6	1153	1 ITAM MOUSE	P05555 mus musculus
3	4358	74.1	1151	2 Q9J130	Q9J130 rattus norv
4	3786.5	64.4	920	2 Q28984	Q28984 sus scrofa
5	3456	58.8	1163	1 ITAD HUMAN	P20702 homo sapien
6	3368	57.6	1162	1 ITAD HUMAN	Q13349 homo sapien
7	3306.5	56.2	1188	2 Q6XAS4	Q6XAS4 mus musculus
8	3306.5	56.2	1188	2 Q6XAS4	Q6XAS4 mus musculus
9	3297.5	56.1	1169	1 ITAX MOUSE	Q9GKH4 mus musculus
10	3225.5	54.9	1161	1 ITAD RAT	Q9GKH4 mus musculus
11	1561	26.6	1165	1 ITAL BOVIN	P61625 bos taurus
12	1539	26.2	1166	2 Q6TVB8	Q6TVB8 bos taurus
13	1539	26.2	1166	2 AAQ90015	AAQ90015 bos tauru
14	1536.5	26.1	1170	1 ITAL HUMAN	P20701 homo sapien
15	1516.5	25.8	1163	1 ITAL MOUSE	P24063 mus musculus
16	1512.5	25.7	1161	2 Q9WTV4	Q9WTV4 mus musculus
17	1502	25.5	1160	2 Q9R200	Q9R200 mus musculus
18	1382	23.5	1196	2 Q9R2F1	Q9R2F1 cyprinus ca
19	1341.5	22.8	1086	2 Q9GHB1	Q9GHB1 homo sapien
20	1329.5	22.6	1187	2 Q9R2F0	Q9R2F0 cyprinus ca
21	1264	21.5	927	2 Q8H2V0	Q8H2V0 cyprinus ca
22	1142.5	19.4	1167	2 Q8H340	Q8H340 bos taurus
23	1136.5	19.3	1167	1 ITAE MOUSE	Q8H340 rattus norv
24	1128	19.2	1179	1 ITAE HUMAN	P38570 homo sapien
25	1097	18.7	1167	2 Q8H341	Q8H341 rattus norv
26	1085.5	18.5	1151	1 ITAL HUMAN	P56199 homo sapien
27	1079.5	18.4	1189	1 ITAH HUMAN	Q9UKX5 homo sapien
28	1071	18.2	1170	1 ITAD BOVIN	P53710 bos taurus
29	1057	18.0	1178	1 ITA2 MOUSE	Q62469 mus musculus
30	1057	18.0	1178	2 Q6P1C7	Q6P1C7 mus musculus
31	1057	18.0	1178	2 AAH65139	AAH65139 mus muscu

ALIGNMENTS

RESULT 1

ID	ITAM HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor M01)			
DE	(Neutrophil adherence receptor).			
GN	Name=ITGAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Artiodactyla; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M01 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

Proc. Natl. Acad. Sci. U.S.A. 86:257-261 (1989).
[6] MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaout M.A.;
RN SEQUENCE OF 1-9 FROM N.A.
RP MEDLINE=92073318; PubMed=1683702;
RX MEDLINE=92144986; PubMed=1346576;
RA Shelley C.S., Arnaout M.A.;
RT "The promoter of the CD11b gene directs myeloid-specific and
RT developmentally regulated expression.";
Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529 (1991).
[7] SEQUENCE OF 1-9 FROM N.A.
RN TISSUE=Blood;
RP MEDLINE=92144986; PubMed=1346576;
RX MEDLINE=92144986; PubMed=1346576;
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;
RT "Characterization of the myeloid-specific CD11b promoter.";
Blood 79:865-870 (1992).
[8] SEQUENCE OF 17-31.
RN MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaout M.A.;
RP MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaout M.A.;
RX MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaout M.A.;
RA Pierce M.W., Remold-O'Donnell E., Todd R.F. III, Arnaout M.A.;
RT "N-terminal sequence of human leukocyte glycoprotein Mol: conservation
RT across species and homology to platelet Iib/IIia.";
Biochim. Biophys. Acta 874:368-371 (1986).
[9] X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
RN MEDLINE=95171458; PubMed=7867070;
RP MEDLINE=95171458; PubMed=7867070;
RX MEDLINE=95171458; PubMed=7867070;
RA Lee J.O., Rieu P., Arnaout M.A., Liddington R.;
RT "Crystal structure of the A domain from the alpha subunit of integrin
RT CR3 (CD11b/CD18).";
Cell 80:631-638 (1995).
[10] X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.
RN MEDLINE=96363671; PubMed=8747460;
RP MEDLINE=96363671; PubMed=8747460;
RX MEDLINE=96363671; PubMed=8747460;
RA Lee J.O., Bankston L.A., Arnaout M.A., Liddington R.C.;
RT "Two conformations of the integrin A-domain (I-domain): a pathway for
RT activation?";
Structure 3:1333-1340 (1995).
[11] X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.
RN MEDLINE=98362595; PubMed=9687375;
RP MEDLINE=98362595; PubMed=9687375;
RX MEDLINE=98362595; PubMed=9687375;
RA Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrichson R.L.,
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;
RT "Cation binding to the integrin CD11b I domain and activation model
RT assessment.";
Structure 6:923-935 (1998).
[12] 3D-STRUCTURE MODELING OF 17-616.
RN MEDLINE=98226734; PubMed=9560195;
RP MEDLINE=98226734; PubMed=9560195;
RX MEDLINE=98226734; PubMed=9560195;
RA Oxvig C., Springer T.A.;
RT "Experimental support for a beta-propeller domain in integrin alpha-
RT subunits and a calcium binding site on its lower surface.";
Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875 (1998).
CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
CC adhesive interactions of monocytes, macrophages and granulocytes
CC as well as in mediating the uptake of complement-coated particles.
CC It is identical with CR-3, the receptor for the iC3b fragment of
CC the third complement component. It probably recognizes the R-G-D
CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
CC of fibrinogen gamma chain.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC granulocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD11b entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03925; AAA59544.1; -
CC EMBL; M18044; AAA59491.1; -
CC EMBL; J04145; AAA59903.1; -
CC EMBL; S52227; AAB24821.1; -
CC EMBL; S52152; AAB24821.1; JOINED.
CC EMBL; S52153; AAB24821.1; JOINED.
CC EMBL; S52154; AAB24821.1; JOINED.
CC EMBL; S52155; AAB24821.1; JOINED.
CC EMBL; S52157; AAB24821.1; JOINED.
CC EMBL; S52159; AAB24821.1; JOINED.
CC EMBL; S52161; AAB24821.1; JOINED.
CC EMBL; S52164; AAB24821.1; JOINED.
CC EMBL; S52165; AAB24821.1; JOINED.
CC EMBL; S52167; AAB24821.1; JOINED.
CC EMBL; S52169; AAB24821.1; JOINED.
CC EMBL; S52170; AAB24821.1; JOINED.
CC EMBL; S52173; AAB24821.1; JOINED.
CC EMBL; S52174; AAB24821.1; JOINED.
CC EMBL; S52180; AAB24821.1; JOINED.
CC EMBL; S52181; AAB24821.1; JOINED.
CC EMBL; S52184; AAB24821.1; JOINED.
CC EMBL; S52189; AAB24821.1; JOINED.
CC EMBL; S52191; AAB24821.1; JOINED.
CC EMBL; S52192; AAB24821.1; JOINED.
CC EMBL; S52203; AAB24821.1; JOINED.
CC EMBL; S52212; AAB24821.1; JOINED.
CC EMBL; S52213; AAB24821.1; JOINED.
CC EMBL; S52216; AAB24821.1; JOINED.
CC EMBL; S52219; AAB24821.1; JOINED.
CC EMBL; S52220; AAB24821.1; JOINED.
CC EMBL; S52221; AAB24821.1; JOINED.
CC EMBL; S52222; AAB24821.1; JOINED.
CC EMBL; S52226; AAB24821.1; JOINED.
CC EMBL; M76724; AAA58410.1; -
CC EMBL; M84477; AAA51960.1; -
CC PIR; A31108; RWHU1B.
CC PDB; 1A8X; Model; @=17-1152.
CC PDB; 1BHO; X-ray; 1/2=-.
CC PDB; 1BHQ; X-ray; 1/2=-.
CC PDB; 1IDN; X-ray; 1/2=-.
CC PDB; 1IDM; X-ray; @=140-331.
CC PDB; 1MIU; X-ray; @=143-334.
CC PDB; 1MF7; X-ray; A=137-331.
CC PDB; 1MF7; X-ray; A=144-337.
CC PDB; 1N92; X-ray; A=140-335.
CC PDB; 1NA5; X-ray; A=144-345.
CC Genew; HGNC:6149; ITGAM.
CC MIM; 120980; -
CC GO; GO:0008305; C:integrin complex; TAS.
CC GO; GO:0007155; P:cell adhesion; TAS.
CC GO; GO:0007043; Integrin_alpha.
CC InterPro; IPR000413; Integrin_alpha.
CC InterPro; IPR002035; VWFA.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; Integrin_alpha; 1.
CC Pfam; PF00092; VWFA; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SM00191; Int_alpha; 5.
CC SMART; SM00327; VWFA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS0234; VWFA; 1.
CC 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
CC Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;
CC Transmembrane.

FT	SIGNAL	1	16
FT	CHAIN	17	1152
		Integrin alpha-M.	
	Query Match	99.1%;	Score 5823.5; DB 1; Length 1152;
	Best Local Similarity	98.8%;	Pred. No. 0;
	Matches 1123;	Conservative	8; Mismatches 5; Indels 1; Gaps 1
QY	1	FNLDTENAMTFOENARGFGQS	VVLQGSVVVVGAPQEIVAA
Dd	17	FNLDTENAMTFOENARGFGQS	VVLQGSVVVVGAPQEIVAA
QY	61	RLOQPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVTVKGLCF	FGSNLRQQPQK 120
Dd	77	RLOQPVAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVTVKGLCF	FGSNLRQQPQK 136
QY	121	FPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKELVSTIMBOLAKSKTLFSLMQSYSEF	180
Dd	137	FPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKEFVTVMBOLAKSKTLFSLMQSYSEF	196
QY	181	RIHFTFKFQNPNRPSLIKPITQLGRTHRTATGLRKVVRELFNITNGARKNAFKILPLL	240
Dd	197	RIHFTFKFQNPNRPSLVKPIITQLGRTHRTATGIKRVVRELFNITNGARKNAFKILVVI	256
QY	241	TGGEKFGLDPLGVIEDVIPDLREGVIRYVLFGDAFRSEKSRQELNTVASKPRDHVFOAN	300
Dd	257	TGGEKFGLDPLGVIEDVIPDLREGVIRYVIGVDFAFRSEKSRQELNTIASKPRDHVFCVN	316
QY	301	NFEALKTIQONLREKI	FAIEGTQTGSSSFHEMSQEGFSAAITSNGPLLTSTVGSYDWAG 360
Dd	317	NFEALKTIQONLREKI	FAIEGTQTGSSSFHEMSQEGFSAAITSNGPLLTSTVGSYDWAG 376
QY	361	GVELYTSKEKSTFFINMTVRVDSMDNDAYLGAAAAIILNRVOSLVILGAPRYOHIGLVAMFR	420
Dd	377	GVELYTSKEKSTFFINMTVRVDSMDNDAYLGAAAAIILNRVOSLVILGAPRYOHIGLVAMFR	436
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGSTDLVLIGAPHYYEOTRGGOVSVCPL	480
Dd	437	QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGSTDLVLIGAPHYYEOTRGGOVSVCPL	496
QY	481	PRGORARWQCDAVLYGEGQGPGWRFGAALTVLGDVNGDKLTDVAIGA	PGBEDNRGAVALVF 540
Dd	497	PRG-RARWQCDAVLYGEGQGPGWRFGAALTVLGDVNGDKLTDVAIGA	PGBEDNRGAVALVF 555
QY	541	HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLT	VGAQHVVLLRSQ 600
Dd	556	HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLT	VGAQHVVLLRSQ 615
QY	601	PVLRVKAIMEFNPREVARNVFECNQDVVKGKEAGEVRCVCLHVOKS	TRDREREQIQSVVT 660
Dd	616	PVLRVKAIMEFNPREVARNVFECNQDVVKGKEAGEVRCVCLHVOKS	TRDREREQIQSVVT 675
QY	661	YDIALDSGRPHSAVFNETKNSTRROTQVILGLTQTCETLKQLPNCIEDPVSPIVLRINF	720
Dd	676	YDIALDSGRPHSAVFNETKNSTRROTQVILGLTQTCETLKQLPNCIEDPVSPIVLRINF	735
QY	721	SLVGTPLSAFGNLRPVLAEDAQRLEFALPFEXKNCNDNI	CODDLSITFSFMGLDCILVWG 780
Dd	736	SLVGTPLSAFGNLRPVLAEDAQRLEFALPFEXKNCNDNI	CODDLSITFSFMGLDCILVWG 795
QY	781	GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASTEV	840
Dd	796	GPREFNVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASTEV	855
QY	841	SGALKSTCSINHIPIPENISEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF	900
Dd	856	SGALKSTCSINHIPIPENISEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF	915
QY	901	QLELPVKVAYVMVTVSHGYSTKYLNFTASENTSRVNMHQHYQVSNLQORSIPTSLVFLVPV	960
Dd	916	QLELPVKVAYVMVTVSHGYSTKYLNFTASENTSRVNMHQHYQVSNLQORSIPTSLVFLVPV	975
QY	961	RLNQTVIWDSPQVTFSENISLSTCHTKERLPSSHDFLAELRKAPVNVCSITAVCORLQCDTP	1020


```

QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAVILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKTQICAGYFASCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
Db 437 ENFGTWEPHTSIKSGIQSGYFASCSVDMADGNTNLILGAPHYEQTRGGQVSVCP 496
QY 481 PRGORARQCDAVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWQCEALLHGDGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYIF 555
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAAGHVLILRSQ 600
Db 556 YGASIASLSASHSHRIIAGHSPGLQYFGQSLSGGKOLTMDGLVDLVAAGHVLILRAQ 615
QY 601 PVLRVKALMEPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
Db 616 PVLREATMEFSPKNLARSVPACQEVVKNKADAGEVRVCLQVKNKTRDRLEGOIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 720
Db 676 YDLALDGRSVVRAPFDETKNGILRRIRVFGTLQKCEETLKLILPCDVNSVPIILRLNY 735
QY 721 SLVGTPLSAFCNLRPVLAEADAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLRPVLAMDAQRFFTAFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREENVTVRNDGDSYQVTFPPPLDLSVRKUSTLQNRQSQRWRL-ACESASSTE 839
Db 796 GPQDFNMSVLRNDGDSYQVTFVYVPSGLSVRKDSASQNPPLTKPFVFKPAESSSSE 855
QY 840 VSGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGNKLLKALVNTSENNMRINKTE 899
Db 856 GHGALKSTTWINHPIPPANSEVTENITFDVDSHAFSGNKLLKALVNTSENNMRINKTE 915
QY 900 FQLELPVKAVYVNTSHGVSTKYLNTFASENTSRVQMHOYQVSNLQSLPISLVFLVP 959
Db 916 FQLELPVKAVYVNTSHGVSTKYLNTFASENTSRVQMHOYQVSNLQSLPISLVFLVP 975
QY 960 VRLNQIVDRPQVTSSENLSSCHTKERLPSHDSFLAELRKAPVNCVSIACQRIQCDI 1019
Db 976 VQLNNVTVMHQPVSQNLSSACHTEQSPHNSFRDQLERTPVLNCSVAVCKRIQCDL 1035
QY 1020 PFFGIQEEFNATLKNLSDWYIKTSHNHLIVSTAEILFNDVSFVILLPGQAFVRSQTE 1079
Db 1036 PSFNTQEIENVTLKGNLSDWYIKTSHGHLILVSSTEILFNDVSFVILLPGQAFVRSQTE 1095
QY 1080 TKVEPEVNPPLIVGSSVGGLLALITALYKLGFFKRYQKMMWSEGGPPGAPQ 1137
Db 1096 TKVEPEVNPPLIVGSSVGGLLALITALYKLGFFKRYQKMMWSEGGPPGAPQ 1153

RESULT 3
Q9J130
ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerria K. Jr.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DDAJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AF268593; AAF81280.1; -.
DR HSSP; F11215; 1BHQ.

```

```

DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 2-A.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02034; VWFA; 1.
DR Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.1%; Score 4358; DB 2; Length 1151;
Best Local Similarity 72.3%; Pred. No. 8.1e-281;
Matches 822; Conservative 151; Mismatches 162; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTNRCDEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPOK 120
Db 77 FLQVPEAVNMSLGLSLAATTVPQLLACGPTVHQCKENTYVNGLCYLFSGNLLRPQ 136
QY 121 FPEALGCCQEDSDIAFLIDGSSIIIPHFRMKELVSTIMEQLKXKTLFSLMYSEEF 180
Db 137 FPEALGCCQCESNIAFLIDGSSINTIDFKMKFEVSTVMDQFKSKTLFSLMYQSD 196
QY 181 RIHFTKEBFQNNPNSLIKPIITOLLGRTHATGLRKVVRRELNTINGARKNAKFLFL 240
Db 197 RIHFTENDPKRNPDPKSHVRPIQLNGRTKASGRKVRRELFOKINGARDNAKILVVI 256
QY 241 TDGEKFGDPLGYEDVPIELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLNVEDVPIEAEEAGIIRYVIGVGNFHKPSRRRELDIASKPAGDHVFQVD 316
QY 301 NFEALKTVOQUREKIFALEGTQSGSSSEFHEMSQEGFSAATISNGPLISTVGSYDWAG 360
Db 317 NFEALNTIRNLOEKIFALEGTQSGTSSFEHMSQEGFSAATISNGPLSGVSGFDWAG 376
QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAVILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKTQICAGYFASCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
Db 437 ENFGTWEPHTSIKSGIQSGYFASCSVDMADGNTNLILGAPHYEQTRGGQVSVCP 496
QY 481 PRGORARQCDAVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWQCEALLHGDGHPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYIF 555
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAAGHVLILRSQ 600
Db 556 YGASIASLSASHSHRIIAGHSPGLQYFGQSLSGGKOLTMDGLVDLVAAGHVLILRAQ 615
QY 601 PVLRVKALMEPNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
Db 616 PVLREATMEFSPKNLARSVPACQEVVKNKADAGEVRVCLQVKNKTRDRLEGOIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 720
Db 676 YDLALDGRSVVRAPFDETKNGILRRIRVFGTLQKCEETLKLILPCDVNSVPIILRLNY 735
QY 721 SLVGTPLSAFCNLRPVLAEADAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLRPVLAMDAQRFFTAFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795

```

781 GRPEFNVTVNRNDGEDSYRTQVTFPPDLISYRKUSTLONQORSQSWRLACESASSTEV 840
 796 DSRDFDVSTLRNDGEDSYGTVKTCYPPSGLSYRKVSASQNFQSKXKFWRIAE-PSSSEG 854
 841 SGALKSTCSINHPFIPENSEVFNTFDVDSKASLGNKLLKXANTVSENMMPTNKTEF 900
 855 QGVLSKSTMDINHPFIPANSEVFNTFDVDSASLGNKLLKXANTVSENMMPTNKTEF 914
 901 QLELPVKYAVYVMTSHGVSTKYLNFASSENTSRVMQHOYQVSNLQORSPLISLFLVPV 960
 915 QLELPVKYAVYVMTSHGVSTKYLNFASSENTSRVMQHOYQVSNLQORSPLISLFLVPV 974
 961 RLMOVTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVACQRIQCDIP 1020
 975 QINKVTIWDPPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVACQRIQCDIP 1034
 1021 FPGIOEFNATLGNLSDFMVYKTSNHLNLLIVSTAEILFNDSPFTLLPGQCAFVRSOTET 1080
 1035 SFNSKEIFNTLQNLFLFDVYIETSHDHLNLLIVSTAEILFNDSPFTLLPGQCAFVRSOTET 1094
 1081 KVEPFEVNPFLPLVGVSSVGLLALLITAAALYKLGFFKQYKDMSEGPPGABPQ 1137
 1095 KVEPFEVNPFLPLVGVSSVGLLALLITAAALYKLGFFKQYKDMSEGPPGABPQ 1151

RESULT 4
 Q28984 PRELIMINARY; PRT; 920 AA.

AC Q28984;
 DT 01-NOV-1996 (TremBLrel. 01, Created)
 DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE CD11b (Fragment)
 GN Name=CD11b;
 OS Sus scrofa (Pig)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Lee J.-K., Schook L.B., Rutherford M.S.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 DR EMBL; U40072; AAB16869.1; -
 DR GO; GO:0008305; C:integrin complex; IEA.
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 4.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS50234; VWF_A; 1.
 KW Cell adhesion; Integrin; Transmembrane.
 FT NON_TER 1
 FT TER 920
 FT SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;

Query Match 64.4%; Score 3786.5; DB 2; Length 920;
 Best Local Similarity 78.4%; Pred. No. 6.4e-243;
 Matches 722; Conservative 85; Mismatches 113; Indels 1; Gaps 1;

118 POKFPALRGCPQEDSDIAFLDGGSGIIPHPFRMKELVSTIMEQLKKSKTLFSLMQYS 177
 1 POKFPALRGCPQEDSDIAFLDGGSGINFLDFQRMKEFVSTVMGQFQSKTLFALMQYS 60
 178 EEPRIHTFKEFQFNPNRSLKIPITQLGRTHATGLRKVRVRELFNITNGAKNFKIL 237
 61 EDYTHFTFDFKRNPSKLLVRPIQLGRTHATGIRKVRVRELFHSGAGRENALKIL 120

238 FLITDGEKFGDPLGYEDVIPLEDRGVIRYVILGFGDAPRSEKSRQBLNVTASKPPRHDVF 297
 121 VITDGEKFGDPLGYEDVIPLEDRGVIRYVILGFGDAPRSEKSRQBLNVTASKPPRHDVF 180
 298 QANNFALKTVQNLREKIFAEGTGTGSSSSPEHEMSOEGFSAAITSGNPILLSTVGSYD 357
 181 QVNNFEAVTQNLQOEKFAEGTGTGSSPECEMSOEGFSAAITSGNPILLSTVGSYD 240
 358 WAGGVLYTSKESKSTINMTRVDSMDNDAYLGAYAAAIILRNRYQSLVGLAPRYQHILGLVA 417
 241 WAGGALHMPKRVIFINTRVDSMDNDAYLGAYAVEVILENQASLVGLAPRYQHTGLV 300
 418 MFRQNTGMESNANVKTQIGAYFGASLSCVVDVDSMGSTDLVILGAPHYEQTRGGQSV 477
 301 MFKQNSGAWKNAIDKGSQISYFGASLSCVVDVDSMGSTDLVILGAPHYEQTRGGQSV 360
 478 CPLPRQARWQCDVILYGEQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAV 537
 361 CPLPQG-RAKMQCRVILCGEQHPGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAV 419
 538 YLPHGTSGSIGSPSHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLL 597
 420 YLPHGTSELGSPSHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLMDLAVGAQGHVLL 479
 598 RSQPVLRVKAIMEFNPREVARNVPECNDQVVKKEAGEVRVCLHVQKSTRDLREGQIQS 657
 480 RSQPVLRVKAIMEFNPREVARNVPECNDQVVKKEAGEVRVCLHVQKSTRDLREGQIQS 539
 658 VVTYDLALDSGRPHSRVAFNETKNTSTROTQVGLTQTCETLKLQNPCEIDPVSPIVLR 717
 540 IITVDLALDQGRPHPRVAFNETKNTSTROTQVGLTQTCETLKLQNPCEIDPVSPIVLR 599
 718 LNFSLVGTPLSAFGLNPLVLAEDAQRLFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCL 777
 600 LNFSLVGTPLSAFGLNPLVLAEDAQRLFTALFPPEKNCNGNDNICQDDLSITFSFMSLDCL 659
 778 VVGGRFNFVTVTVNRNDGEDSYRTQVTFPPDLISYRKUSTLONQORSQSWRLACESAS 837
 660 VVGGRFNFVTVTVNRNDGEDSYRTQVTFPPDLISYRKUSTLONQORSQSWRLACESAS 719
 838 TEVSALKSTCSINHPFIPENSEVFNTFDVDSKASLGNKLLKXANTVSENMMPTNKTEF 897
 720 TEVSALKSTCSINHPFIPENSEVFNTFDVDSKASLGNKLLKXANTVSENMMPTNKTEF 779
 898 TEPQLELPVKYAVYVMTSHGVSTKYLNFASSENTSRVMQHOYQVSNLQORSPLISLFLVP 957
 780 TEPQLELPVKYAVYVMTSHGVSTKYLNFASSENTSRVMQHOYQVSNLQORSPLISLFLVP 839
 958 VPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVACQRIQ 1017
 840 VPVRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCIAVACQRIQ 899
 1018 DIPFGIOEFNATLGNLSDFMVYKTSNHLNLLIVSTAEILFNDSPFTLLPGQCAFVRSOTET 1038
 900 DIPFGIOEFNATLGNLSDFMVYKTSNHLNLLIVSTAEILFNDSPFTLLPGQCAFVRSOTET 920

RESULT 5
 ITAX_HUMAN
 ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
 AC P20702; Q81VAG;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 05-JUL-2004 (Rel. 44, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
 alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
 GN Name=ITGAX; Synonyms=CD11c;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=8816645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of a
 RL leukocyte adhesion glycoprotein, p150,95.";
 RN EMBO J. 6:4023-4028(1987).
 RP [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
 RT molecule";
 RL J. Biol. Chem. 265:2782-2788(1990).
 RN [3]
 RP ERRATUM.
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751(1990).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young J.W., Green E.D., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Kryzanski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RL and p150,95 leukocyte adhesion proteins.";
 RL J. Immunol. 138:2381-2383(1987).
 CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 CC granulocytes.
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81695; AAA51620.1; -.
 DR EMBL; M29165; -; NOT_ANNOTATED_CDS.
 DR EMBL; M29487; AAA51620.1; ALT_SEQ.
 DR EMBL; M29482; AAA51620.1; JOINED.
 DR EMBL; M29483; AAA51620.1; JOINED.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29485; AAA51620.1; JOINED.
 DR EMBL; M29486; AAA51620.1; JOINED.
 DR EMBL; BC038237; AAA51620.1; JOINED.
 DR PIR; A36584; RWHUIC.
 DR PDB; IN3Y; X-ray; A=141-338.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; -.
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0009887; P:organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWFA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWFA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 19
 FT CHAIN 20 1163 Integrin alpha-X.
 FT DOMAIN 20 1107 Extracellular (Potential).
 FT TRANSMEM 1108 1128 Potential.
 FT DOMAIN 1129 1163 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 351 FG-GAP 2.
 FT DOMAIN 165 351 VWFA.
 FT REPEAT ? 401 FG-GAP 3.
 FT REPEAT 402 453 FG-GAP 4.
 FT REPEAT 455 517 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 466 474 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1131 1135 GPFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 655 712 By similarity.
 FT DISULFID 771 777 By similarity.
 FT DISULFID 848 863 By similarity.
 FT DISULFID 998 1022 By similarity.
 FT DISULFID 1027 1032 By similarity.
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).
 FT VARIANT 48 48 W -> R (in dbSNP:11574633).
 FT /FTID=VAR_018672.
 FT T -> S (in Ref. 4).
 FT T -> A (in Ref. 4).
 FT T -> S (in Ref. 4).
 FT G -> A (in Ref. 2).
 FT E -> K (in Ref. 4).
 FT D -> L (in Ref. 4).
 FT I -> V (in Ref. 4).
 FT SEK -> TPHPQDNV (in Ref. 4).
 FT STRAND 1161 1163
 FT 150 157

```
FT TURN 160 161
FT HELIX 164 178
FT TURN 179 180
FT TURN 183 185
FT STRAND 186 193
FT STRAND 197 201
FT HELIX 203 208
FT HELIX 212 216
FT TURN 217 218
FT STRAND 226 226
FT HELIX 228 236
FT TURN 237 240
FT HELIX 242 244
FT TURN 245 245
FT TURN 248 249
FT STRAND 251 258
FT STRAND 263 263
FT HELIX 269 278
FT TURN 279 280
FT STRAND 282 288
FT HELIX 290 293
FT TURN 296 297
FT HELIX 298 304
FT HELIX 310 312
FT STRAND 313 316
FT HELIX 319 325
FT HELIX 326 334
FT TURN 335 335
SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5DB8F CRC64;

Query Match 58.8%; Score 3456; DB 1; Length 1163;
Best Local Similarity 60.6%; Pred. No. 9,3e-221;
Matches 684; Conservative 143; Mismatches 296; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFQSQVVLQGRVVVVGAPQEIIVAAANQGRSLYQCDYSTGSCBPI 60
DB 20 FNLDTIELTAFRVDAGSDSVQVYANSVVVVGAPQKITAANQGTGGLYQCGYSTGACEPI 79

QY 61 RLQVPEAVNMSLGLSLAATTPPOLLACGPTVHTCSENTYVKGCLFLFGSNLQOQOK 120
DB 80 GLQVPEAVNMSLGLSLASTTSPQLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137

QY 121 FPEARLGCPOEDIAFLDGGSIIPHPDFRMKELVSTIMEOLKSKTFLSLMOYSSEF 180
DB 138 LPVSRQECPRQCDIVFLDGGSSISRNFATMNFVRAVISQFORPSTQFSLMQFSNKF 197

QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKKVVRLEFNITNGARKNAKILFLL 240
DB 198 QTHFTFEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLFGDAPRSEKSRQELNTVASKPRDHFVQAN 300
DB 258 TDGKKEGSDLYKDVPMDAAGIIRYAIQVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

QY 301 NFEALTKVONQREKIFATEGTQTSSSSFEHEMSQEGFSAATISNGPLLSITVGSYDWAG 360
DB 318 DFDALNDIQNKKEKIFALEGHTETSSSFEHEMSQEGFSAATISNGPLLSITVGSYDWAG 377

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILNRRVQSVILGAPRYQHIGLVAMPR 420
DB 378 GAFLYPPNMSPTFINMSQENVDNRDSYLGYSTELALWKGVSILVLGAPRYQHTGKAVIET 437

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGTDVLVLGAPHYVEOTRGQSVQVCP 480
DB 438 QVSRQWRMAEAVTGTQISYFGASLCSDVDNDSGTDVLVLGAPHYVEOTRGQSVQVCP 497

QY 481 PRGORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 498 PRGWR-RWWCDVLYGEQHPWGRFGAALTVLGDVNGDKLTDVVGAPGEENRGAVYLF 556

QY 541 HGTSGGSGPSHSQRAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTVGAQHVLIRSQ 600
DB 557 HGVLPSPISPSHSQRAGSKLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
```

```
QY 601 PVLVRKAIMBEPNPREVARNVFECNDQVYVKGKEAGEVVRVCLHVQKSTDRLEREGIOISVVT 660
DB 617 PVLWVGSMQFIPAEIPRSAPFECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676

QY 661 YDLALDSGRPHSRAVFNETKNSTRQTVGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 LDIALDPGRLSPRATFOETKNRSLSRVRVLGLKAHCENFNLLLESCVEDSVTPTILRLNF 736

QY 721 SLVCTPLSARGNLRPLVLAEDAQRLLFTALFPFCKNCGNDNICQDDLSITFFSMDCLVVG 780
DB 737 TLVGKPLLAFLARNLRPLAADAQRYFTASLPFCKNCGADHICQDNLGIFSFPGLKSLVVG 796

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSTYRKVSTLQNRQSRQSRWRLACESASSTEV 840
DB 797 SNLENAEAVWVNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHLTCDSPAPVG-- 854

QY 841 SGALKTSCTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 900
DB 855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVLGDRLLTANVSSENTPRTSKTTF 914

QY 901 QLELPVKYAVVMVVTSHGVSTKYNFTAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 959
DB 915 QLELPVKYAVTVVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSINFVWP 974

QY 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNCISIAVCQRIQCDI 1019
DB 975 VELNQEAVMVDVEVSHPCNSLRCSEKIAAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034

QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAETLFDNSVFTLAPGGQAFVRSOTE 1079
DB 1035 PSFSVQBELDFTLKNLSFGWVRQLQKKVSVSWAEITFDTSVYSQPLPGQAFMRAQT 1094

QY 1080 TKVREFEVPNPLPLIVGSSVGGLLILALITAAALYKLGFFKQYKDMMESE 1128
DB 1095 TVLEKIKYHNPTPLIVGSSIGGLLLILALITAVLYKVGFFKQYKEMMEE 1143

RESULT 6
ITAD HUMAN
ID Q13349; Q15575; Q15576; PRT; 1162 AA.
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Splice;
RX MEDLINE=961111956; PubMed=8777714;
RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690(1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sp1 and Sp3.";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
RT beta 2-integrin alpha subunit.";
```

RL Gene 171:291-294 (1996).
 RN [4]
 RP INTERACTION WITH VCAM1.
 RX MEDLINE=99059842; PubMed=9841932;
 RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;
 RT "alpha2beta2 integrin is expressed on human eosinophils and functions
 as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
 1).";
 RL J. Exp. Med. 188:2187-2191 (1998).
 RN [5]
 RP INTERACTION WITH VCAM1.
 RX MEDLINE=99370002; PubMed=1043935;
 RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazquez R., Hoffman P.A.,
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
 binding interface between I domain and VCAM-1";
 RL J. Immunol. 163:1984-1990 (1999).
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
 VCAM1. May play a role in the atherosclerotic process such as
 clearing lipoproteins from plaques and in phagocytosis of blood-
 borne pathogens, particulate matter, and senescent erythrocytes
 from the blood.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
 associates with beta-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
 lines and subsets of peripheral blood leukocytes and strongly on
 tissue-specialized cells, including macrophages foam cells within
 atherosclerotic plaques, and on splenic red pulp macrophages.
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL; U37028; AAB38547.1; -
 DR EMBL; U40274; AAB60634.1; -
 DR EMBL; U40275; AAB60635.1; -
 DR EMBL; U40276; AAB60636.1; -
 DR EMBL; U40277; AAB60637.1; -
 DR EMBL; U40279; AAB60638.1; -
 DR EMBL; U40278; AAB60638.1; JOINED.
 DR EMBL; AF187881; AAF62875.1; -
 DR HSSP; P11215; IBHQ.
 DR Genew; HGNC:6146; ITGAD.
 DR MTM; 602453; -
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0016337; P:cell-cell adhesion; TAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00327; VWA; 1.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 17 potential.
 FT

FT CHAIN 18 1162 Integrin alpha-D.
 FT DOMAIN 18 1100 Extracellular (Potential).
 FT DOMAIN 1101 1121 Potential.
 FT DOMAIN 1122 1162 Cytoplasmic (Potential).
 FT REPEAT 32 85 FG-GAP 1.
 FT REPEAT 86 7 FG-GAP 2.
 FT DOMAIN 150 332 VWFA.
 FT REPEAT 350 400 FG-GAP 3.
 FT REPEAT 401 452 FG-GAP 4.
 FT REPEAT 454 516 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 465 473 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1127 1131 GFFKR motif.
 FT DISULFID 67 74 By similarity.
 FT DISULFID 106 124 By similarity.
 FT DISULFID 655 710 By similarity.
 FT DISULFID 769 775 By similarity.
 FT DISULFID 846 861 By similarity.
 FT DISULFID 994 1018 By similarity.
 FT DISULFID 1023 1028 By similarity.
 FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 500 500 Missing (in Ref. 2).
 FT CONFLICT 515 518 GHPW -> ATP (in Ref. 2).
 FT CONFLICT 825 825 L -> V (in Ref. 2).
 FT CONFLICT 984 984 V -> A (in Ref. 2).
 SQ SEQUENCE 1162 AA; 126885 MW; F296AIA3545SD77D CRC64;
 Query Match 57.6%; Score 3388; DB 1; Length 1162;
 Best Local Similarity 58.9%; Pred. No. 3.1e-216;
 Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFOENARGFGSVVQLQGSRRVVGAPQBIIVAAQSGSLVYQCDYSTGSCPI 60
 DB 18 FNLDVEEPTIFQBDAGFGSVVQFGSRLVVGAPLEVAANOTGRLYDCAATGMCPQI 77
 QY 61 RLQVPVAVNNSLGLSLAATTSPQLACGPTVHOTCSENTYVKGCLFLGSLNRQPOK 120
 DB 78 PLHIRPEAVNNSLGLTLAASTNGSRLACGPTLHRVCGENSYSKSGCLLGSRW-ELIQT 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMRKELVSTIMEQLKSKTLFSLMQYSEEF 180
 DB 137 VPDATPECPHQEMDIVFLIDGSGSIQDNDFNQMKFVQAVMGQFEGDTLFLALMQYSNLL 196
 QY 181 RIHFTFEQNNENPESLKIPIQLLGRTHATGLRKKVRELNITNGARKNAFKLFL 240
 DB 197 KIHFTFQRTSPSQSLVDPIVLQKGLTFTATGILTITVVTQLFHHKNGARKSAKILIVI 256
 QY 241 TDGEKFGDPLGVEDVPIPELDREGVIRYVLGFGDAPRSEKSRQELNTVASKPPDDHVFQAN 300
 DB 257 TDQKYKDPLEYSVDVPIQAEKAGIIRYAGVGHAFQGPARTARQELNTISSAPDQHVFKVD 316
 QY 301 NFALKTVQNQRKIFAIEGTQTGSSSFHEMSQEGFSAITSNGPLISTVGSYDWAG 360
 DB 317 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMQGLFLGAVGSFWSG 376
 QY 361 GVFLYTSKESSTFINMTRVDSMDNAYLGYAAAIILNRNVSQSLVGLAPRYQHTGLVAMFR 420
 DB 377 GAFLYPPNMSPTFINNSQENVDMDRSDYLGSTELALWKGQNLVGLAPRYQHTGKAVIF 436
 QY 421 QNTGMWSENANVKGTOIGAYFGASLCSDVDVDSNGSDTLVLIGAPHYYEOTRGQNSVCL 480
 DB 437 QVSRQWRKKAETVGTQIGSYFGASLCSDVDSDGSDTLVLIGAPHYYEOTRGQNSVCL 496

RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
 RA Tsuchiya H.;
 RT "Isolation of genes selectively expressed by dendritic cells.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC !- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis (By similarity).
 CC !- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2 (By similarity).
 CC !- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC !- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.
 CC !- SIMILARITY: Contains 1 VWFA domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: AF211864; AAF23492.1; --
 DR HSSP: P20702; IN3Y.
 DR MGD; MGI:96609; Itgax.
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.
 DR InterPro: IPR000413; Integrin_alpha.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF01839; FG-GAP; 3.
 DR Pfam: PF00357; Integrin_alpha; 1.
 DR Pfam: PF00092; VWF; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWF_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 19 Potential.
 FT CHAIN 20 1169 Integrin alpha-X.
 FT DOMAIN 20 1116 Extracellular (Potential).
 FT TRANSMEM 1117 1137 Potential.
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 7 FG-GAP 2.
 FT DOMAIN 152 330 VWFA.
 FT REPEAT 402 402 FG-GAP 3.
 FT REPEAT 403 454 FG-GAP 4.
 FT REPEAT 456 518 FG-GAP 5.
 FT REPEAT 519 577 FG-GAP 6.
 FT REPEAT 582 634 FG-GAP 7.
 FT CA_BIND 467 475 Potential.
 FT CA_BIND 531 539 Potential.
 FT CA_BIND 594 602 Potential.
 FT SITE 1140 1144 GFFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 656 711 By similarity.
 FT DISULFID 770 776 By similarity.
 FT DISULFID 858 873 By similarity.
 FT DISULFID 1007 1031 By similarity.
 FT DISULFID 1036 1041 By similarity.
 FT CARBOHYD 89 89 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 267 267 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 393 393 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 734 734 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 949 949 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 1059 1059 N-linked (GlcNAc. .) (potential).
 FT CARBOHYD 1084 1084 N-linked (GlcNAc. .) (potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;
 Query Match 56.1%; Score 3297.5; DB 1; Length 1169;
 Best Local Similarity 56.2%; Pred No. 3.3e-210;
 Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;
 QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVVGAPQRIEVAANQGRSLVOCYSTGSCBPI 60
 DB 20 FNLDAAEKLTHPHMDGAEFGHSLVQYDSSVWVVGAPKEIKATNQIGGLYKCGYHTGNCBPI 79
 QY 61 RLQVPEAVNMSLGLSLAATSPPLLACGPTVHTCTSENTVVKGLGCLFGLNLRQOPK 120
 DB 80 SLQVPEAVNISLGLSLAATNPSSLACGPTVHTCTSENTVVKGLGCLFGLNLRQOPK 138
 QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKKSKTLFSLMOYSERF 180
 DB 139 FTAQOECPKQDDIVFLIDGSGSISSDTEKMLDFVKAQMSQLQSPSTRFSLMQFSDYF 198
 QY 181 RIHFTKEFQNNPNRSLIKPITQLLGRTHATGRKVVRELFNTNGARKNAKILFL 240
 DB 199 RVHFTFNNFISTSPSLSLGVRQLRGVYTTASAKHVITELFTTQSGARQDATKVLIV 258
 QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFSEKSRQELNNTVASKPRDHVFOAN 300
 DB 259 TDKRGQGNLSYDSVIPMAEASIIIRYAIQVGFNEHSHKQELKAIASMPHEVYVSV 318
 QY 301 NFEALKTVONQLREKIFAIEGTQSSSFEHMQEFGSAATISNGPLSTVGSYDWAG 360
 DB 319 NFDALKDINQJKEKIFAIEGTETSSSTFELEMQEFGSAVTPDGPVLGAVGSPWSG 378
 QY 361 GVFLYTSKEKSTFINMTRVDSDDNDAYLGYAAAILRNVRQSLVLGAPYQHIGLVAMPR 420
 DB 379 GAFLYPSNMRPTFINMQENEDMDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVI 438
 QY 421 QNTGWESNANVKGQIGAYFGASLCSDVDVDSNGSDTLVLGAPHYVETRGQGVSCPL 480
 DB 439 QESRHWPKSEVRGTQIGSYFGASLCSDVMDRGGSDTLVLGAPHYVETRGQGVSCPM 498
 QY 481 PRGQARWOCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEENRGAYLF 540
 DB 499 P-GVGRWECGTTLHGEQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAYIF 557
 QY 541 HGTSGSIGSPHSQRIAGSKLSPRLQYFGQSLSGGGQDLTMDGLVDLTVAQGHVLLRSQ 600
 DB 558 HGASRQDIAPSPSORISASQIPRIQYFGQSLSGGGQDLTMDGLVDLAVSGKRVLLRTR 617
 QY 601 PVLRVKAIEMFPREVARNVFECNDQVVKGEAGEVRVCLHVCKSTRDLREGIOISVVT 660
 DB 618 PILRVSPVTHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIESPTQL--GDLRSTVI 675
 QY 661 YDLALDSGRPHRAVENETKNSRRTQVIGLTQTCTELKQLPNCIEDPVSPVILRLNF 720
 DB 676 FDLALDHGRLSTRAIEKTKTRALTAVKILGNKHCSVKLLLPACVEDSVTPIRLNF 735
 QY 721 SLVGTPLSAFNGRLPVLAEADAQLFTALPFFPKNGNDNICQDDLITITFSMLDCLVVG 780
 DB 736 SLVGVPISSILQNLQPLAVDDQYFTASLPFFKNGADHICQDDLVSFVGFPPDLKLTVG 795
 QY 781 GPREFNVTVVRNDGEDSVRTQVTFPFDLDSYRKVSTLQ-----NORSORSWR 829
 DB 796 SDLELVNDVTVSNDGEDSVGTVTTLFPVGLSFRRAVGQVFLRKEDQWQRGGQSLH 855
 QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
 DB 856 LMCD--STPDRSQGLNSTSCSRHVIIFRGSSQMTFLVTFVSPKAEGLDRLLRARVGE 913
 QY 890 NNMPRNKTEFOLELPVKVAVYVMVTVSHGVSTKYLNFTASE-NTSRVWQHYQVSNLQOR 948
 DB 914 NNVEGTPKTTFOLELPVKVAVYVMVTVSHGVSTKYLNFTASE-NTSRVWQHYQVSNLQOR 973
 QY 949 SLPLSLVFLVPLVRLNQTWIDVRPQVTFSENLSCTCHKRLPSHSDPLAELRKAPVNC 1008
 DB 974 DVPVSINFVWPIELKGEAVW-TVMVSHQPQPLTQCYNRLKPTQFDLLTHMQKSPVLD 1032

```

QY 1009 IAVCORIOCDIPFFQIOEEFNATKGNLSFDWYIKTSHNHLIVSTAILENDSVFTLLP 1068
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1033 IADCHLHRCIDPSLGLDELFIPLKGNLSFGWISLQKLVLLSEAITNTSVYSQLP 1092
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1069 GQGFVRSQTEKTVKFFVFPNPLPLIVGSSVGGLLALLIITAALYKLGFFKQYKDMNSE 1128
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1093 GQEAFLRAQTKVLEMYKHPVPLVGVSSVGGLLALLIITAILYKAGFFKQYKEMLEE 1152
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 10
ID ITAD RAT STANDARD; PRT; 1161 AA.
AC Q90VE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.W., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
EA "Cloning of rat alpha D, a novel beta 2 integrin.";
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RL VCAM1. Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC -! FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -! SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -! DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -! SIMILARITY: Belongs to the integrin alpha chain family.
CC -! SIMILARITY: Contains 7 FG-GAP repeats.
CC -! SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWFA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR KEGG; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potentially.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.

```

```

FT REPEAT 88 ? FG-GAP 2.
FT DOMAIN 152 334 VWFA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GFFKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 54.9%; Score 3225.5; DB 1; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.1e-205;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;

QY 2 NLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQSGSLVQCDYSTGSCPIR 61
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 NLDVEEPIVFEADAASFGQTVVQFGSRLVVGAPLEAVAVNQTGLYDCAATGWCQIV 80
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 62 LQVPEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTVYKGLCFGLNLRQOPQKF 121
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 LRSPLAVNMSLGLSLVATNNAQLLACGPTAQRACVKNMYAKGSCLLIGSSL-QFIQAV 139
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 122 PEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSEFR 181
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 PASMPECPROEMDI AFLIDGSGSINQDRFAQMDKDFVXALMGFEFASTTLFSLMOYSNLK 199
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 182 IHFTKFEQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNTNGARKNAKILFLIT 241
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 200 THFTTFEKNILDPSLVDPIVQLQGLTYATGIRTWEEELFHSKNGSRKSAKILLVIT 259
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 242 DGEKFGDPLGYEDVIPELDREGVTRYVILGFGDAPRSEKSRQELNVTASKPRDHVFQANN 301
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 DGQYRDPLEYSVDVIPAADKAGIIRYAIGVDAPQEPALKELTIGSAPPQDHVFKVGN 319
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 302 FEALKTQVQNLREKI FALGTOGTGSSSFHEHMSOEGFSAITNSGPLLSTVSGVDRAGG 361
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 320 FAALRSIQRLQEKIFAIEGTQSRSSSFQHEHMSOEGFSSALTSDDPVLGVSFSSWGG 379
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 362 VFLYTSKEKSTFFINWTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMPRQ 421
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 380 AFLYPNTRPTFFINMSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVVITQ 439
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 422 NTGMWESNANVKGTQIGAYFGASLCSVDNSNGSTDVLVIGAPHYEQTRGGQVSVCPPL 481
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDNRDGSDDLVLGAPHYEQTRGGQVSVFPVP 499
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 482 RQQRARMQCDVLYGEOGPWGRFGAALTVLGVNKGKLTVDVAICAGEEDNRGAVLFFH 541
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 500 -GVKRWQCEATLHGEQHPWGRFGVALTVLGVNKGKLTVDVAICAGEEDNRGAVLFFH 558
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 542 GTSGSGISPSHSQRIAGSKLSPLRQYFGQSLSGQDLTMDGLVLTVAQGHVLLLSQP 601
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 559 GASPLEIMPSPSQRTVGSQSLRLQYFGQSLSGQDLTQDGLVLDLAVGAQGHVLLLSLP 618
  ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
QY 602 VLRVKAIMEFNPVARNVFECDQVVKKEAGEVYVCLHVOKSTDRDLREGQIQSVVY 661
D 619 LKVLBSIRFAPMEVAKAYQWERTPTVLEAGEATVCLTVHKGSPDLL--GNVGSVRY 676
QY 662 DLALDSGRPHSAVNETKNSRRQTQVLGLTQTCETLKLQDPNCIEPVPVSLVRLNFS 721
D 677 DLALDPGLISRAIDFENKNCITLGRKTLGLGDHCTGVKLLPDCVEDAVSLIIRLNF 736
QY 722 LVGTPLSAPGNRPLVLAEDAQLFTALPFFKNCNGNDNICQDDLITFSFMSLDCILVGG 781
D 737 LVRDSASP--RNLPVLAAGSDHITASLPFFKNCQELLCEGLGISENFSGQLVVLVGG 795
QY 782 PFEFNVTVTRNDGDSYRTQVTFPPLDLSYRKVSTLQNRQSRWRLACASASTEV 841
D 796 SPELTVTVTVMWEGEDSYTLVKFYFAGLSYRRVTGTQ-QPHQVPLRLACAEFAAQBD 854
QY 842 GALKTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANTVSENMPRTNKTBF 901
D 855 --LRSSCSINHPIFREGAKTTFMITFDVSYKAFGLDRLLLRKAKASSENKPDTKTAPQ 912
QY 902 LELPKYAVYVMTVSHGVSTKYLNFTASNTGR-VWQHOYQVSNLQGRSLPISLFLVPV 960
D 913 LELPKYVTVYTLISROEDSTNVHFNFSHSGRRQEAHRYRVNLSPLKLVAVNFWVPV 972
QY 961 RLNQIVINDRPOVTFSENLS--TCHTKERLPSHDFLAELKAPVNCVSIACQRIQCDI 1019
D 973 LINGVAVMD---VTLSPPAGQVSCVSKPPQNPDLTQIQRSLVDCSIALCLHFRCDI 1029
QY 1020 PFGIQEENATLKNLSPDWIKTSHNHLIVSTAELFNDVFTLLPGQAFVRSQTE 1079
D 1030 PSLDIQDELDFILRNLGNSFGWVSQTLQEKVLLVSEAEITFDVSVYSQLPQEAFLRAQVE 1089
QY 1080 TKVEFPEVNPPLVGVSGVGLLILALITAILYKLGFPRQYKQNM 1126
D 1090 TTLEEVVYVEPIFLVAGSVGGVGLLILALITVLYKLGFPRQYKQNM 1136

RESULT 11
ITAL_BOVIN STANDARD; PRT; 1165 AA.
AC P61625;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=ITGAL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA Pett T., Zecchinon L., Baise E., Desmecht D.;
RT "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT characterisation and comparison with the human and murine
RT glycoproteins."
RL Gene 325:97-101(2004).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The integrin I-domain (inset) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
```

```
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY267467; AAP94035.1; -.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS02344; VWFA; 1.
CC Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
CC Repeat; Signal; Transmembrane.
CC SIGNAL 1 23 Potential.
CC CHAIN 24 1165 Integrin alpha-L.
CC DOMAIN 24 1084 Extracellular (Potential).
CC TRANSMEM 1085 1105 Potential.
CC DOMAIN 1106 1165 Cytoplasmic (Potential).
CC REPEAT 40 88 FG-GAP 1.
CC REPEAT 89 146 FG-GAP 2.
CC REPEAT 153 324 VWFA.
CC REPEAT 347 397 FG-GAP 3.
CC REPEAT 398 452 FG-GAP 4.
CC REPEAT 454 513 FG-GAP 5.
CC REPEAT 515 572 FG-GAP 6.
CC REPEAT 575 627 FG-GAP 7.
CC CA_BIND 465 473 Potential.
CC CA_BIND 527 535 Potential.
CC CA_BIND 587 595 Potential.
CC SITE 1111 1115 GPFKR motif.
CC DISULFID 71 78 By similarity.
CC DISULFID 108 126 By similarity.
CC DISULFID 650 704 By similarity.
CC DISULFID 768 774 By similarity.
CC DISULFID 842 858 By similarity.
CC DISULFID 994 1009 By similarity.
CC DISULFID 1017 1048 By similarity.
CC CARBOHYD 33 33 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 86 86 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 185 185 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 646 646 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 667 667 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 723 723 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 859 859 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 894 894 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 929 929 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 1056 1056 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 1067 1067 N-linked (GlcNAc...) (Potential).
CC SEQUENCE 1165 AA; 128725 MW; DABBJA3F1E1463CB CRC64;
Query Match 26.6%; Score 1561; DB 1; Length 1165;
Best Local Similarity 34.4%; Pred. No. 1.4e-94;
Matches 407; Conservative 200; Mismatches 468; Indels 108; Gaps 33;
QY 1 FNLDTENAMTFQENARG--FGQSVVQLQGSVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCE 58
D 24 YNLDVRHVQNFQFFPLAGRHFGRVRLQV--GNGVVVGAPSE---GNSMGNYQQCPETGDC 79
QY 59 PIRLQVFEAVNNSLGLSLAATSPPELLACGPTVHQTGSENTYVKGLCFGLFSGNLROOP 118
D 80 PVTLS--SNYTSKYLQMTLATDPTSDNLACDPGLSRTCDQNYLSGLCYLIHENLRGPV 137
QY 119 QKFPALRCQPOEDSDIAFLIDGSGSIIPHDFRKMELVSTIMEQLKSKTLPFLMQYSE 178
D 138 LQCHPGYQECIKGNVDLVFLFDGNSLQODEFEKIVDFMKDVMKLSNSYSQFAAVQFST 197
QY 179 ERIHPTKEFQNNPNRSLIKPITOLLGRTHATGLRKVVRLEFNI TWGARKNAFKILF 238
D 198 YRTEFTFLDIYRQDDALLAGVHKMRLTLTFTFGAINVAVEVFPDLGARPDATKVL 257
```



```
QY 652 EQIOSVVTDIALDSGRPHSAVNETKNSTRROTUVLGLTQTCETLKLQLPNCIEDPV 711
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
662 QGHLVANLTYTLQLDGHRTRSGLPFGGKHLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 712 SPIVLRLNPSL---VGTPLS--AFGNLRPVLAEDAQRFTALFPPEKNCNNDNICODDLS 766
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 721 SPINVSLSYSLWEERGTTPDRALDRDIPILKPSHLETKETIPPEKNGCEDKNCADLK 780
QY 767 ITFSFMSLDCLVVGGRPREFNVTVRNDCGESYRTQVTFFFLDLISYRKVSTLQNRQSR 826
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSPFQGLSFRKVEIL---KPHS 837
QY 827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANV 886
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 838 HVPVGCCELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKSGWDFIELQANV 896
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 887 TS-----ENNMPRTNKTETQLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMOH 938
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 897 SCNNEDSSILLEDSNATS-----IPVMYPINVLTKQOENSTLYISFTPKSPRIHHVKH 949
QY 939 QYQV-----SNLQORSLPISLFLVPLRNLQTVI---WD---RPQVTFSS-ENLSSTCHTKE 987
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMBPPVNCSPRNLESPSDEAE 1008
QY 988 RLPSSHDFLAELRKAPVNCSTAVCQRIODIPFGIQEENFATLKGNSLFDWYIKTSHN 1047
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1009 -----SCSFGT--EPRCPIDF---ROEILVQVNGMVELRGTTIKAS-S 1044
QY 1048 HLLIVSTABILFNDSVFTLLPGGAFVRSGTETKVEPFEVPNPPLPIVINGSSVGGLLILLAL 1107
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1045 MSLCSSLAISFNSSKHFLHGNASM-AQVVMKVDLVYKEKMLYLVLSGIGLLLLFL 1103
QY 1108 ITAALYKLGFFKRYQKDDM-----SEGPPGABPQ 1137
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1104 IFITALYKVGFFKRNLEKMEANVDASSEIPGEDAGQPELEKE 1145
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
AAQ90015 PRELIMINARY; PRT; 1166 AA.
ID AAQ90015, AC AAQ90015, DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RA SEQUENCE FROM N.A.
RP Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a."
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY382558; AAQ90015.1;
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.2%; Score 1539; DB 2; Length 1166;
Best Local Similarity 34.3%; Pred. No. 4e-93;
Matches 405; Conservative 199; Mismatches 472; Indels 106; Gaps 33;

QY 1 FNLDTENAMTFQENARG--FGQSVQLQSGRVVVGAPQELIVANQSGLYQCDYSGTSC 58
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 25 YNLDRVHQVNFSPFLAGRHFGYVLQV-GNGVYVVGAPSE---GNSMGNYQQCPETGDL 80
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 PIRLOVPVEAVNMSLGLSLAATTSPPOLLAGCGTTHVTCENTYVVKGLCLFLGNSLRQOP 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 81 PVTLSS--SNYTSKYLGNLATDFTSDNLLACDGLSRTCDQNIYLSGLCYLIHENLRGPV 138
QY 119 QKPEALRGCPQEDSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSE 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 14
ITAL_HUMAN

alpha-subunit and its expression in COS cells."; J. Immunol. 147:369-374(1991).
 [2]
 SEQUENCE OF 24-42.
 MEDLINE=85188276; PubMed=3887182;
 Springer T.A., Teplov D.B., Dreyer W.J.;
 "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 Nature 314:540-542(1985).
 -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
 ICAM3 and ICAM4. Is involved in a variety of immune phenomena
 including leukocyte-endothelial cell interaction, cytotoxic T-cell
 mediated killing, and antibody dependent killing by granulocytes
 and monocytes. Mice expressing a null mutation of the alpha-L
 subunit gene demonstrate impaired tumor rejection and impaired
 leukocytes recruitment.
 -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
 associates with beta-2.
 -!- SUBCELLULAR LOCATION: Type I membrane protein.
 -!- TISSUE SPECIFICITY: Leukocytes.
 -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 with I-domains do not undergo protease cleavage.
 -!- SIMILARITY: Belongs to the integrin alpha chain family.
 -!- SIMILARITY: Contains 7 FG-GAP repeats.
 -!- SIMILARITY: Contains 1 VWFA domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M60778; AAA39426.1; --
 PIR; I56126; I56126.
 HSP; P20701; IDGQ.
 MGP; MG1:96606; Itgal.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR02035; VWF_A.
 Pfam; PF01839; FG-GAP; 1.
 Pfam; PF00357; Integrin_alpha; 1.
 Pfam; PF00092; VWFA; 1.
 PRINTS; PR01185; INTEGRINA.
 PRINTS; PR00453; VWFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWFA; 1.
 SMART; SM00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; VWFA; 1.
 Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;
 Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.
 SIGNAL 23
 CHAIN 24 1163 Integrin alpha-L.
 DOMAIN 24 1087 Extracellular (Potential).
 TRANSMEM 1088 1108 Potential.
 DOMAIN 1109 1163 Cytoplasmic (Potential).
 REPEAT 39 88 FG-GAP 1.
 REPEAT 89 146 FG-GAP 2.
 DOMAIN 153 325 VWFA.
 REPEAT 348 398 FG-GAP 3.
 REPEAT 399 454 FG-GAP 4.
 REPEAT 455 514 FG-GAP 5.
 REPEAT 516 573 FG-GAP 6.
 REPEAT 576 628 FG-GAP 7.
 CA_BIND 466 474 Potential.
 CA_BIND 528 536 Potential.
 CA_BIND 588 596 Potential.
 SITE 1111 1115 GFFKR motif.
 FT 70 77 By similarity.
 DISULFID 108 126 By similarity.
 DISULFID 147 199 By similarity.
 DISULFID 651 705 By similarity.
 DISULFID 767 773 By similarity.

FT	DISULFID	840	856	By similarity.
FT	DISULFID	993	1009	By similarity.
FT	DISULFID	1017	1048	By similarity.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	185	185	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	270	270	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	444	444	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	668	668	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	724	724	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	728	728	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	776	776	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	857	857	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	880	880	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	890	890	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	899	899	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	927	927	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1056	1056	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	1163 AA;	128343 MW;	A7A3078489E8232F CRC64;

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;
 Best Local Similarity 33.9%; Pred. No. 1.2e-91;
 Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

QY	1	ENLDTENAMTFQENA-RGFGQSVVQLQSGRVVVGAPQELVAANQRGSLYOCDSYSGSCEP	59
Db	24	YNLDRTPQSFQAQGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRSTSEFCQP	79
QY	60	IRLQVPVEAVNMSLGLSLAATTSPQLIACGPTVHQTCSENTYVKGCLFLFGSNLRQCPQ	119
Db	80	VSLH-GSNHTSKYLGMTLATDAAGSLIACDPGLSRCTQNTYLSGLCYLPQSLGEML	138
QY	120	KPPALRCQPEDSDIAFLIDGSGSIIPHDPFRMKELVSTIMEQLKSKTKTLFSLMQYSEE	179
Db	139	QNRPAYQECMKGVDLNLFDFGQSQSLDRKDKILEFMKQVWRKLSNTSYQFAAVQFSTD	198
QY	180	FRIHETKEF-QNNPNRSLIKPTQLLGRTHRTATGLRKVVRELFININGARKNAFLIF	238
Db	199	CRTEFTLDYVYKQKNPDVLGSGVQPMFLNTTFRAINVYVAHVFKESGARPATKVLV	258
QY	239	LLTDGEKF--GDPLGYEDVPELDEGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHV	296
Db	259	IITDGEASDKGNISAHD-----ITRIIIGIKHFVSQVQKTLHIIFASEVEBFV	309
QY	297	FOANFEALKTVOQLREKIFAISGTQTGSSSEHEMSQEGFSAATSGNPLSTVGSY	356
Db	310	KILDTFEKLDLFTDLQRIYAIETGNRQDLTSFNMELSSSGISADLSKGHAUVGAVGAK	369
QY	357	DWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGYAA--IILNRVQSLVLGAPRYQHIG	414
Db	370	DWAGFLDLREDLQGFVGOEPLTSDVRGGYLGTYVAMWTSRSRPLAAGAPRYQHV	429
QY	415	LVAMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQRG	472
Db	430	QVLLFQAPAEAGGRWNQTKIEGTQIGSYFGELCSVDLDQDGEALLIGLAPLFFGEORG	489
QY	473	GQSVVCPLPRQARWOCDAVLVYGEQOPMGRFGAALTIVLGDVNGDKLTDVAIGAPGED	532
Db	490	GRVFTY---QRRQSLFEMVSELQDGPYPLGRFGAAITALTIDINGDRITDVAVGAPLEE-	545
QY	533	NRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGGQDLTDGDLVDTVAQG	592
Db	546	-QGAVYIFNGKPG--GLSPQSPQRIQGAQVFGIRFWFGSRHGVKDLGGDLADVVVGAEG	603
QY	593	HVLLRSQPVLRKAIMEFNPREVARNVFECNDQVVKGBAG-EVRYCLHVQKSTRDLR	651
Db	604	RVVLSRPVVDVVTLELSFSPPEIPVHEVECSVSAREQKHGKVKAKAFRIKPLTPQ--P	661
QY	652	EGQIQSVVTVDLALDSGRPHSRVFNFTKXSTRQTQVLGLTQTCETLKLQPLNCIEDPV	711
Db	662	QGRLLANLSTLQLDGHMRMRSGLFPDGSHELSTSTIP-DKSCLDLFFHFHFCIQDLI	720
QY	712	SPIVLRNFSLV---GTPLSAFGN-LRPVLAEADAQRLFTALFFPEKNCNDNTICQDDL	767

Db

721

SPINVSINFSLEBEGTTRDQGRAMQPIILRPSIHTV-TKEIPFEKNCGEDKKCEANLTL

779

QY

768

TFSEMSLDCLVVGGP-----REFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQN

821

Db

780

SSPARS-----GPIRLMSSASLAVEMTSLNSGEDAYWVRDLDFPRGLSPRKVEMLQ-

831

QY

822

QSQRSWRLACESASSTEVSGAL-KSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKL

880

Db

832

--PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQVMTLNSWEDFV

887

QY

881

LILKANVTSEN-NMPRTNKTEFOLELPVKYAVVMVVTSHGVSTKXLYNFTASENTSRVMOHQ

939

Db

888

ELNGTVHCENENSSLOEDNSAATHIPVLPVNVILTYEQENSILYISFTKPGKTOQVQHV

947

QY

940

YQVSNLQGRSLPISLVLPVRLNQTVIWDRPQ----VTFSENLS--TCHTXE-RLP

990

Db

948

YQV-----RIQPSAYDHNMT-LEALVGVPRHSEDLITYTWSVQTDPLVTCHSEDLKRP

1001

QY

991

SHSDFLAELRKAPVNVCSIAVCQRIQCDDIPFGIQEEFNATLKNLSFDWYIKTSHNLL

1050

Db

1002

SSE---AEQPCLEGV-----QFRCPVIF--RWEILIQVTGTVELSKEIKAS-STLS

1046

QY

1051

IVSTAEILFNDSVFTLLPGQAFVRSQTEKVEPPEVNPPLPLIVGSSVGGGLLILALITA

1110

Db

1047

LCSSLSVSNSSKHFLYLSKA-SEAQVLVKVDLHEKEMLHVYVLSGIGGLVLLFLFL

1105

QY

1111

ALYKLGFFKRYQKXMM-SEGPPGAEP

1136

Db

1106

ALYKVGFFKRNKLEKMEADGGVNGSP

1132

Search completed: November 9, 2004, 12:17:27
Job time : 142.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQKDMSEGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5839	99.3	1153	2	AAW65090 Human Bet
2	5839	99.3	1153	3	AAW07360 Human CD1
3	5839	99.3	1153	5	AAU80252 Human int
4	5839	99.3	1153	5	ABG61469 Human Bet
5	5839	99.3	1153	5	AAO14428 Integrin
6	5839	99.3	1153	7	ADD25615 Binding d
7	5829	99.1	1153	2	AAO4136 Alpha sub
8	5823.5	99.1	1152	8	ADM99589 Human int
9	5816.5	98.9	1152	8	ADP12435 Protein e
10	3460	58.9	1163	8	ADP44061 Human CD1
11	3446	58.6	1163	8	ADN02004 Human inf
12	3446	58.6	1163	8	ADQ17510 Human sof
13	3437	58.5	1163	2	AAO7120 p150.95 a
14	3423	58.2	1163	2	AAW65091 Human Bet
15	3423	58.2	1163	3	AAW07361 Human CD1
16	3423	58.2	1163	5	ABG61470 Human Bet
17	3421	58.2	1163	6	ABU07406 Protein d
18	3421	58.2	1163	7	ADG32005 Human hom
19	3388	57.6	1161	2	AAW78166 Human bet
20	3388	57.6	1161	2	AAW23049 Human bet
21	3388	57.6	1161	2	AAW57491 Human bet
22	3388	57.6	1161	2	AAW65089 Human bet
23	3388	57.6	1161	2	AAW72825 Human alp
24	3388	57.6	1161	2	AAW73342 Human alp
25	3388	57.6	1161	3	AAW07359 Human alp

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment: psoriasis;
XX KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KW rheumatoid arthritis.

XX OS Homo sapiens.

XX PN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using immobilised or
labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11b subunit which is
used to describe a method for identifying compounds that modulate the
interaction of the beta-integrin alpha-d subunit with a binding partner
of alpha-d which involves contacting an alpha-d polypeptide with an alpha
-d binding partner, one of which is immobilised and the other of which is
labelled, in the presence of a test compound, and determining if the
compound affects binding between the alpha-d polypeptide and alpha-d
binding partner, where the alpha-d polypeptide is alpha-d or its fragment
comprising the cytoplasmic, transmembrane or extracellular domain of
alpha-d. Compounds that modulate alpha-d binding could be used to treat.

ABG61468 Human Bet
AAW23064 Human Bet
AAW65106 Human Bet
AAW72837 Human alp
AAW73343 Human alp
AAB07376 Human alp
ABG61485 Human Bet
AAW78169 Rat alpha
AAW23062 Rat beta
AAW60004 Rat alpha
AAW72824 Rat alpha
AAB07374 Rat alpha
ABG61483 Rat Beta2
AAW65104 Rat beta-
AAW73345 Rat alpha
AAW23061 Mouse bet
AAW60003 Mouse alp
AAW65103 Mouse bet
AAW72836 Mouse alp
AAW73347 Mouse alp

26 3388 57.6 1161 5 ABG61468
27 3372.5 57.4 1161 2 AAW23064
28 3372.5 57.4 1161 2 AAW65106
29 3372.5 57.4 1161 2 AAW72837
30 3372.5 57.4 1161 2 AAW73343
31 3372.5 57.4 1161 3 AAB07376
32 3372.5 57.4 1161 5 ABG61485
33 3215.5 54.7 1161 2 AAW78169
34 3213.5 54.7 1161 2 AAW23062
35 3213.5 54.7 1161 2 AAW60004
36 3213.5 54.7 1161 2 AAW72824
37 3213.5 54.7 1161 3 AAB07374
38 3213.5 54.7 1161 5 ABG61483
39 3206.5 54.5 1161 2 AAW65104
40 3206.5 54.5 1161 2 AAW73345
41 3201 54.4 1161 2 AAW23061
42 3201 54.4 1161 2 AAW60003
43 3201 54.4 1161 2 AAW65103
44 3201 54.4 1161 2 AAW72836
45 3201 54.4 1161 2 AAW73347

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSQVVQLQGRVVVVGAPQEIIVAAANQKSLQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFQSQVVQLQGRVVVVGAPQEIIVAAANQKSLQCDYSTGSCBPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRLQOQPK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRLQOQPK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTKLFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTKLFLSLMOYSEEF 196
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILV 256
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVLGFGDAPRSEKSRQELMTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGVEDVPELDRGVIRYVLGFGDAPRSEKSRQELMTVASKPRDHVFOAN 316
QY 301 NFEALKTVQNLQREKIFALEGTQTGSSSFEHMQEGFSAITGNGPLLSLVGSDWAG 360
DB 317 NFEALKTVQNLQREKIFALEGTQTGSSSFEHMQEGFSAITGNGPLLSLVGSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCLP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCLP 496
QY 481 PRGQARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQOPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLLRSQ 600
DB 557 HTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLLRSQ 616
QY 601 PVLRYKAIMERNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRSGOIQSVVT 660
DB 617 PVLRYKAIMERNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRSGOIQSVVT 676
QY 661 YDLALDSGRPHRAVENETKNSRRTOTQVLGLTQTCETLKLQPLNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHRAVENETKNSRRTOTQVLGLTQTCETLKLQPLNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNLRLPVLADDAQLFTALFPFEKNCNDNICDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNLRLPVLADDAQLFTALFPFEKNCNDNICDDLSITFSFMSLDCLVVG 796
QY 781 GREFNVTTVNRDGEDSVRTQVTFPPFDLSVRKYSTLQNGRSQSRWLACESASSTEV 840
DB 797 GREFNVTTVNRDGEDSVRTQVTFPPFDLSVRKYSTLQNGRSQSRWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 857 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916
QY 901 QLELPVKYAVYMWVTHSGVSTKYLNFTASENTSRVMQHVQVSNLQCRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYMWVTHSGVSTKYLNFTASENTSRVMQHVQVSNLQCRSLPISLVFLVPV 976

QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020
DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1036
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGQGAFAVRSQTET 1080
DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGQGAFAVRSQTET 1096
QY 1081 KVEPPEVENPPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSSGGPPGAPQ 1137
DB 1097 KVEPPEVENPPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSSGGPPGAPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX AC AAB07360;

XX DT 17-JAN-2001 (first entry)

XX DE Human CD11b protein sequence.

XX KW Human; macrophage infiltration inhibition; alpha_d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11b.

XX OS Homo sapiens.

XX PN WO200029446-A1.

XX PD 25-MAY-2000.

XX PF 16-NOV-1999; 99WO-US027139.

XX PR 16-NOV-1998; 98US-00193043.

XX PR 08-JUL-1999; 99US-00350259.

XX PA (ICOS-) ICOS CORP.

XX PI Gallatin MW, Van Der Vieren M;

XX XX WPI; 2000-387751/33.

DR Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.

XX Example 5; Fig 1; 270pp; English.

XX Integrins are a class of membrane-associated molecules that participate
CC in cellular adhesion. Integrins are made up of an alpha subunit and a
CC beta subunit. One class of human integrins are restricted to expression
CC in white blood cells and have a common beta2 subunit: the leukocyte
CC integrins, leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
CC have an important role in immune and inflammatory responses. The present
CC protein sequence is the human integrin alpha subunit CD11b. This sequence
CC was used in an alignment to identify a novel beta2 integrin alpha
CC subunit: alpha d (AAA60014 and AAB07359). The present sequence has
CC approximately 60% identity to the protein sequence of alpha d. The
CC alpha d gene and protein may be useful in therapy for diseases linked to
CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the
CC inhibition of macrophage infiltration at the site of a central nervous
CC system injury. The monoclonal antibodies can also be used to detect and
CC diagnose Crohn's disease

SQL Sequence 1153 AA;
Query Match 99.3%; Score 5839; DB 3; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGGQSVVQLQSGSRVVVGAPEIIVAAQNRGLSLQCDYSGSCPEI 60
DB 17 FNLDTENAMTFOENARGGQSVVQLQSGSRVVVGAPEIIVAAQNRGLSLQCDYSGSCPEI 76
QY 61 RLQVPVEAVNMLGLSLAATSPPOLLAGCTVTHQTCSENTYVVKGLCFGLFSGNLRQOPK 120
DB 77 RLQVPVEAVNMLGLSLAATSPPOLLAGCTVTHQTCSENTYVVKGLCFGLFSGNLRQOPK 136
QY 121 FPEALRGCPQSDIAFLIDGSGSTIIPHDFRMKELVSTIMEOLKSKTILFSLMOYSEEF 180
DB 137 FPEALRGCPQSDIAFLIDGSGSTIIPHDFRMKELVSTIMEOLKSKTILFSLMOYSEEF 196
QY 181 RIHFTFEFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILLFLL 240
DB 197 RIHFTFEFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILLVVI 256
QY 241 TDGKFGDPLGYEDVIELDREGVIRYVGLGDAFRSKSRQELNTVASKPRDRHVFQAN 300
DB 257 TDGKFGDPLGYEDVIELDREGVIRYVGLGDAFRSKSRQELNTVASKPRDRHVFQAN 316
QY 301 NFEALKTVONOLREKIPIAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTVONOLREKIPIAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVRVSDMNDVILGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRVSDMNDVILGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR 436
QY 421 QNTGWMESNANVKGTQIAGYFASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL 480
DB 437 QNTGWMESNANVKGTQIAGYFASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL 496
QY 481 PRGQARWQCDVILYGGQGPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNREGAVYLF 540
DB 497 PRGQARWQCDVILYGGQGPWGFAGALTIVLGVNGDKLTDVAIGAPGEDNREGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQLYFGSLSGQDLTMDGLVDLTWGAQGHVLLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQLYFGSLSGQDLTMDGLVDLTWGAQGHVLLLRSQ 616
QY 601 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
DB 617 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSPIVILNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPVSPIVILNF 736
QY 721 SLVGTPLSAFGRNLRPLVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFFGMSLDCLVWG 780
DB 737 SLVGTPLSAFGRNLRPLVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFFGMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGDSYQTVTFPPDLISVRKVTSLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGDSYQTVTFPPDLISVRKVTSLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANTVSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQVQVSNLGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMQHQVQVSNLGRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1036

QY 1021 PFGIOEBFNATLKGNSLDFDWIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1080
DB 1037 PFGIOEBFNATLKGNSLDFDWIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSQTE 1096
QY 1081 KVEPFEVNPPLIVGSSVGLLLALITALYKLGPEKROYKDMWSGGPPGAPQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGLLLALITALYKLGPEKROYKDMWSGGPPGAPQ 1153
RESULT 3
AAU80252
ID AAU80252 standard; protein; 1153 AA.
XX AAU80252;
AC AAU80252;
XX 15-JUL-2002 (first entry)
XX Human integrin 1 alpha-M subunit protein.
XX Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;
KW inflammatory disease; autoimmune disorder; Crohn's disease;
KW human immunodeficiency virus; HIV; myocardial infarction;
KW Sjorgen's syndrome; rheumatoid arthritis.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 499..500
XX /note="Encoded by GGG CAG AGG"
XX WO200218583-A2.
XX 07-MAR-2002.
XX 31-AUG-2001; 2001WO-US027227.
XX 01-SEP-2000; 2000US-0229700P.
XX (BLOO-) CENT BLOOD RES INC.
XX Springer TA, Shimoaka M, Lu C;
XX WPI; 2002-382964/41.
XX N-PSDB; ABK50046.
XX Modified integrin-I or integrin I-like domain polypeptide useful as an
XX immunogen to produce antibodies specific to polypeptide, comprises a
XX disulfide bond such that polypeptide is stabilized in a desired
XX conformation.
XX Disclosure; Page 109-112; 112pp; English.
XX This invention relates to a modified integrin-I or integrin I-like domain
XX polypeptide comprising at least one disulfide bond so that the domain is
XX stabilised in a desired conformation. The polypeptide of the invention
XX may have antiinflammatory or immunosuppressive activities. The
XX polypeptides of the invention have an open conformation and are useful as
XX immunogens to produce antibodies that selectively bind to integrin I-
XX domain; and for identifying a modulator of integrin activity, or of
XX interaction of an integrin and a cognate ligand. The polypeptide of the
XX invention, or antibodies (preferably anti-LFA-1 antibody) is useful for
XX treating or preventing an integrin mediated disorder which is an
XX inflammatory or autoimmune disorder in a subject and for inhibiting the
XX binding of an integrin to a cognate ligand such as Crohn's disease,
XX nephritis; human immunodeficiency virus (HIV), myocardial infarction,
XX Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic
XX composition comprising the peptide of the invention is useful for
XX treating an integrin mediated disorder in a subject. The polypeptides
XX of integrin-mediated disorders are useful as reagents for diagnosis
XX and/or active or antigenic fragments are useful as reagents for diagnosis
XX of integrin-mediated disorders. The present sequence represents the human
XX integrin-1 alpha-M protein subunit used to generate the mutant
XX polypeptides of the invention

XX	SQ	Sequence 1153 AA;
		Query Match 99.3%; Score 5839; DB 5; Length 1153; Best Local Similarity 98.9%; Pred. No. 0; Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY	1	FNLDTENAMTFOENARGFGVQSVOLOQSRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI 60
DB	17	FNLDTENAMTFOENARGFGVQSVOLOQSRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI 76
QY	61	RLOVPVEAVNMISLGLSIAATTPSPQLACGTVHTQTCSENTYVKGLCFLFSGNLRQQPQK 120
DB	77	RLOVPVEAVNMISLGLSIAATTPSPQLACGTVHTQTCSENTYVKGLCFLFSGNLRQQPQK 136
QY	121	FPEALRGCPOSDIAFLIDSGSIIPHDFPRMKELYSTIMEOLKKSKTLFSLMOYSSEF 180
DB	137	FPEALRGCPOSDIAFLIDSGSIIPHDFPRMKEFVSTVMEQLKKSKTLFSLMOYSSEF 196
QY	181	RIHFTFEFQNPNPRSIIKITOLLGRTHATGLRKVKVRELFINITNGARKNAFKILFL 240
DB	197	RIHFTFEFQNPNPRSIVKIITOLLGRTHATGIKRVKVELFINITNGARKNAFKILVVI 256
QY	241	TGGEKFGDPLGYEDVIPELDREGVIRYVLGPDAPFRSKSQEINLTWASKPPRDHVFOAN 300
DB	257	TGGEKFGDPLGYEDVIPEDAREGVIRYVIGDAPFRSKSQEINTLIASKPPRDHVFOVN 316
QY	301	NFEALKTVONQLREKIFAIETGTOTGSSSFSEHEMSQEGFSAAITSNGPLSTVGSYDWAG 360
DB	317	NFEALKTIQNLREKIFAIEGTQTGSSSFSEHEMSQEGFSAAITSNGPLSTVGSYDWAG 376
QY	361	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPYOHIGLVAMFR 420
DB	377	GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPYOHIGLVAMER 436
QY	421	QNTGMWESNANYKGTOIGAYFGASLCSDVDVNSGSTDLVLIGABHYEBQTRGGQVSVCPL 480
DB	437	QNTGMWESNANYKGTOIGAYFGASLCSDVDVNSGSTDLVLIGABHYEBQTRGGQVSVCPL 496
QY	481	PRGQPARWCDAVLYGEOGPWGREGAALTIVLDVNGDKLTVAIGAPGEDNKGAVLYF 540
DB	497	PRGQPARWCDAVLYGEOGPWGREGAALTIVLDVNGDKLTVAIGAPGEDNKGAVLYF 556
QY	541	HGTSGSGISPSSHQRISAKLSPLRYQGSLSGQDLTMDELVDLTVGAQHVLRLRSQ 600
DB	557	HGTSGSGISPSSHQRISAKLSPLRYQGSLSGQDLTMDELVDLTVGAQHVLRLRSQ 616
QY	601	PVLRVKAIMERNPREVARNVFECNDQVVKGEAGBRVCLHVQKSTRDRLBEGIQSVVT 660
DB	617	PVLRVKAIMERNPREVARNVFECNDQVVKGEAGBRVCLHVQKSTRDRLBEGIQSVVT 676
QY	661	YDLALDSGRPSRAVFNETKSNRTQTVLGLTOTCETLKLPNCIEDPVSPVLRINF 720
DB	677	YDLALDSGRPSRAVFNETKSNRTQTVLGLTOTCETLKLPNCIEDPVSPVLRINF 736
QY	721	SLVGTPLSAFGNLRPVLAEDAQRLETFALFPPEKNCNDNICODDLSITFFSMISDCLVVG 780
DB	737	SLVGTPLSAFGNLRPVLAEDAQRLETFALFPPEKNCNDNICODDLSITFFSMISDCLVVG 796
QY	781	GPREFNVTVVRNDEGDSYRTQVTPFFPLDSLRSKVSTLQNRQSRWRLACESASTEV 840
DB	797	GPREFNVTVVRNDEGDSYRTQVTPFFPLDSLRSKVSTLQNRQSRWRLACESASTEV 856
QY	841	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASIGNKLKLLKANVTSENMPRNKTKEF 900
DB	857	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASIGNKLKLLKANVTSENMPRNKTKEF 916
QY	901	QLELPKYAVVMVTSHGVSITYLNFTASENTSRVMHQYQVSNLQSRSLPSLFLVPV 960
DB	917	QLELPKYAVVMVTSHGVSITYLNFTASENTSRVMHQYQVSNLQSRSLPSLFLVPV 976
QY	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEURKAPVNVNCSTAVCQRIQCDDIP 1020

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFQSQVVLQGGSRVVVGAPOEIVAAQORSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTQENARGFQSQVVLQGGSRVVVGAPOEIVAAQORSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNNLSGLSLAATTSPPOLLACGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 120
Db 77 RLQVPVEAVNNLSGLSLAATTSPPOLLACGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 136

QY 121 FPEALRGCPQDSDIATLIDGSGSTIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQDSDIATLIDGSGSTIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFEFQNNPNSRLIKPITQLLGRTHATGLRKVVRELENTINGARKNAFKILPLL 240
Db 197 RIHFTFEFQNNPNSRLIKPITQLLGRTHATGLRKVVRELENTINGARKNAFKILVVI 256

QY 241 TDGKFGDPLGYEDVPELDRGVIRYVGLGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQVN 316

QY 301 NFEALKTQONLRREKIFAETGTQSGSSSFHEMSQEGFSAITNSGPELLSTVGSYDWAG 360
Db 317 NFEALKTQONLRREKIFAETGTQSGSSSFHEMSQEGFSAITNSGPELLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGWESNANVKQTQI GAVFGASLCSDVDVDSNGSTDLVLIGAPHYEQTGQGVSCPL 480
Db 437 QNTGWESNANVKQTQI GAVFGASLCSDVDVDSNGSTDLVLIGAPHYEQTGQGVSCPL 496

QY 481 PRGORARQWQCDVLYGQGGPWGFGAALTVLGVDNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGORARQWQCDVLYGQGGPWGFGAALTVLGVDNGDKLTDVAIGAPGEDNREGAVYLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDQLTMDGLVDLTGVAQGHVLLLRQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGQDQLTMDGLVDLTGVAQGHVLLLRQ 616

QY 601 PVLRVKAIMBFPREAVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLRVKAIMBFPREAVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMDCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYQTVPFPDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYQTVPFPDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANTVSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANTVSENMPRTNKTEF 916

QY 901 QLELPVKYAVYVMVTSHGVSSTKYNLFTASNTSRVMQHQYQVSNLQORSPLISLVLFPV 960
Db 917 QLELPVKYAVYVMVTSHGVSSTKYNLFTASNTSRVMQHQYQVSNLQORSPLISLVLFPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISVACQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISVACQRIQCDIP 1036

QY 1021 PFGIQEENATLKGNSLSPDWIKTSHNHLILVSTAEILFNDSVFTLLPQGAFAVSQET 1080
Db 1037 PFGIQEENATLKGNSLSPDWIKTSHNHLILVSTAEILFNDSVFTLLPQGAFAVSQET 1096

QY 1081 KVEPEFVENPPLVGVSSVGGLLLLALTAALYKLGFEKROYKDMMSGGPPGAEPQ 1137
Db 1097 KVEPEFVENPPLVGVSSVGGLLLLALTAALYKLGFEKROYKDMMSGGPPGAEPQ 1153

RESULT 5

AAO14428
ID AAO14428 standard; protein; 1153 AA.

XX AAO14428;

AC AAO14428;

DT 03-MAY-2002 (first entry)

XX Integrin Mac-1 alpha subunit.

DE Mac-1; integrin alpha subunit; variant integrin inserted domain protein;
KW open conformation; integrin related inflammatory disorder;
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;
KW reperfusion; hypovolemic shock; infarction; cerebral shock;
KW viral infection; cancer; gene therapy; vaccine;
KW bioactive agent screening.

XX Unidentified.

OS WO200204521-A2.

PN 17-JAN-2002.

PD 09-JUL-2001; 2001WO-US021805.

XX 07-JUL-2000; 2000US-0216600P.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX (BLOO-) CENT BLOOD RES.

XX Springer T;

XX WPI; 2002-148167/19.

XX New integrin I domain protein having alteration in at least 2
noncontiguous regions and exits in an open conformation, useful for
treating, preventing or suppressing inflammatory or immunological
disorders.

XX Example 1; Fig 1F; 90pp; English.

XX The invention comprises structurally biased variant integrin inserted (I)
domain proteins, wherein the alterations to the protein occur in at least
two noncontiguous regions. Specifically the variant integrin I domain
proteins are structurally biased to exist in the open conformation,
thereby altering the binding ability of the protein. The invention also
comprises nucleic acids encoding the variant integrin I domain proteins.
The integrin I domain proteins and nucleic acids are useful for treating,
preventing or suppressing integrin related inflammatory and immunological
disorders (e.g. rheumatoid arthritis). The variant integrin I domain
proteins and nucleic acids can also be used for treating: ischaemia/
reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral
infection; and cancer. The variant integrin I domain nucleic acids and
proteins may be used in gene therapy, as vaccines and to screen for
bioactive agents. The present amino acid sequence represents the Mac-1
alpha subunit of integrin

XX Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSIVOLQGSRRVVGAPQEIIVAAANQORSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTQENARGFGQSIVOLQGSRRVVGAPQEIIVAAANQORSLYQCDYSTGSCBPI 76
QY 61 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFSLNLRQQPK 120
DB 77 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFSLNLRQQPK 136
QY 121 FPEARLGCPEQSDIAFLIDSGSIIIPDFFRMKELVSTIMEQLKSKTSLFSLMOYSBEF 180
DB 137 FPEARLGCPEQSDIAFLIDSGSIIIPDFFRMKELVSTIMEQLKSKTSLFSLMOYSBEF 196
QY 181 RIHFTFKFQNNPNSRLKPIITOLLGRTHATGLKVVRELFNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNSRLKPIITOLLGRTHATGLKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVLGDAFRSEKSRQELNTVASKPPRDHVFQAN 316
QY 301 NFEALKTQVONQUREKI PAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDRAG 360
DB 317 NFEALKTQVONQUREKI PAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDRAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496
QY 481 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGVNGDKLTVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HTSGSGTSPSHSQRISAGSLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600
DB 557 HTSGSGTSPSHSQRISAGSLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616
QY 601 PVLVRKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGLRQIQSVVT 660
DB 617 PVLVRKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGLRQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETIKLQPLNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETIKLQPLNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQLFTALFPFKNKCGNDNICDDLSITFSFMSLCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAQLFTALFPFKNKCGNDNICDDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDSYRTQVTPFPPLDLVSRKVSTLQNRQSRWSRLACESASTEV 840
DB 797 GPREFNVTVVRNDSYRTQVTPFPPLDLVSRKVSTLQNRQSRWSRLACESASTEV 856
QY 841 SGALKTSKCSINHPIPENSEVFNITFDVDSKASIGNKLLKANVTSENPNRTNKTTF 900
DB 857 SGALKTSKCSINHPIPENSEVFNITFDVDSKASIGNKLLKANVTSENPNRTNKTTF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYNFTASENTSRVQHQYQVSNLQGRSLPISLVLPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYNFTASENTSRVQHQYQVSNLQGRSLPISLVLPV 976
QY 961 RLNQTVWRDPQVTFSENISSTCHTKERLPSPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020
DB 977 RLNQTVWRDPQVTFSENISSTCHTKERLPSPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1036
QY 1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQTET 1080
DB 1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVEPFVEVFNPLPLVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPFGAEPQ 1137

DB 1097 KVEPFVEVFNPLPLVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPFGAEPQ 1153
RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX
AC ADD25615;
XX
DT 15-JAN-2004 (first entry)
XX
Binding domain-immunoglobulin fusion protein-associated protein #85.
XX
DE Binding domain; immunoglobulin; fusion protein; cytostatic;
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;
KW neuroprotective; hinge region; immunoglobulin heavy chain;
KW CH2 constant region; CH3 constant region; IgG1;
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.
XX
OS Unidentified.
XX
XX US2003118592-A1.
PN
XX
PD 26-JUN-2003.
XX
XX 25-JUL-2002; 2002US-00207655.
PF
XX
PR 17-JAN-2001; 2001US-0367350P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX
PA (GENE-) GENE-CRAFT INC.
XX
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
PI WPI; 2003-801317/75.
XX
DR
XX
XX
PT
PT
PT
XX
PS
XX
CC The invention relates to a binding domain-immunoglobulin fusion protein comprising a binding domain polypeptide that is fused to an immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain CH2 constant region polypeptide that is fused to the hinge region polypeptide, and an immunoglobulin heavy chain CH3 constant region polypeptide that is fused to the CH2 constant region polypeptide. The hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues; where the mutated human IgG1 immunoglobulin hinge region polypeptide contains 2 cysteine residues, where the first cysteine is not mutated; a mutated human IgG1 immunoglobulin hinge region polypeptide, derived from (a) having 3 or more cysteine residues, where the mutated human IgG1 immunoglobulin hinge region polypeptide contains no cysteine residues. The binding domain-immunoglobulin fusion protein is capable of at least one immunological activity comprising antibody dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The binding domain polypeptide is capable of specifically binding to an antigen. Also included are an isolated polynucleotide encoding the binding domain-immunoglobulin fusion protein, a recombinant expression construct comprising the polynucleotide (operably linked to a promoter), a host cell transformed or transfected with a recombinant expression construct, producing the binding domain-immunoglobulin fusion protein, a

pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder. The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

XX Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 7; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVLQGSRRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOQK 136

QY 121 FPBALRGCPEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
Db 137 FPEALRGCPEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196

QY 181 RIHFTKEFQNNPRSLIPITQILGRTHATGLRKVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTKEFQNNPRSLIPITQILGRTHATGLRKVRELFNITNGARKNAFKILV 256

QY 241 TDEKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHVFQAN 300
Db 257 TDEKFGDPLGYEDVPELDREGVIRVVLGFGDAFRSEKSRQELNTIASKPRDHVFQVN 316

QY 301 NFPAKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATNSGPFLLSTVGSYDWAG 360
Db 317 NFPAKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATNSGPFLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTINMTRVDSMDNDAYLGAAAILLNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTINMTRVDSMDNDAYLGAAAILLNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL 496

QY 481 PRQORARWQCDVLYGEQGPWGRFGAALTDLVGDVNGDKLTDVAIGAPGEEDNRGAYL 540
Db 497 PRQORARWQCDVLYGEQGPWGRFGAALTDLVGDVNGDKLTDVAIGAPGEEDNRGAYL 556

QY 541 HGTSGSGISPHSQRTAGSKLSPRLQYFGSLGSGQDLTMDGLVDLTVGAGQHVLRLRSQ 600
Db 557 HGTSGSGISPHSQRTAGSKLSPRLQYFGSLGSGQDLTMDGLVDLTVGAGQHVLRLRSQ 616

QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTFDRREGQIOSVVT 660
Db 617 PVLVRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTFDRREGQIOSVVT 676

QY 661 YDLALDSGRPHSAVFNETKNSTRQTQVLGLTQTCETLKLQPNCLIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSAVFNETKNSTRQTQVLGLTQTCETLKLQPNCLIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEDAQRILFTALFPPEKXGNDNICQDDLSTITFSFMSLDCI 780
Db 737 SLVGTPLSAFGNLRPVLAEDAQRILFTALFPPEKXGNDNICQDDLSTITFSFMSLDCI 796

QY 781 GPREENVTVRNDGEDSYRTQVTFPPDLVSKVKSTLQNSORSQSWRLACESASSTEV 840
Db 797 GPREENVTVRNDGEDSYRTQVTFPPDLVSKVKSTLQNSQRSQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916

QY 901 QLELPVKYAYVMVTSHGVSSTKYLNFASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAYVMVTSHGVSSTKYLNFASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFESNLSSTCHTKERLPSHSDLAELRKAPVWVNCISIAVCORIQC 1020
Db 977 RLNQTVIWDPRQVTFESNLSSTCHTKERLPSHSDLAELRKAPVWVNCISIAVCORIQC 1036

QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET 1096

QY 1081 KVEPFVFPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKXDMMEGGPPGAPQ 1137
Db 1097 KVEPFVFPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKXDMMEGGPPGAPQ 1153

RESULT 7
AAR041136
ID AAR041136 standard; protein; 1153 AA.

XX AAR041136;
AC AC
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 07-SEP-1990 (first entry)
XX
DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW non-specific defence system; integrin gene superfamily.
XX
OS Synthetic.

Key	Location/Qualifiers
Region	1..116
Modified-site	/label= signal_peptide
Modified-site	86..88
Modified-site	/label= putative N-glycosylation site
Modified-site	240..242
Modified-site	/label= putative N-glycosylation site
Modified-site	391..393
Modified-site	/label= putative N-glycosylation site
Modified-site	469..471
Modified-site	/label= putative N-glycosylation site
Modified-site	693..695
Modified-site	/label= putative N-glycosylation site
Modified-site	697..699
Modified-site	/label= putative N-glycosylation site
Modified-site	735..737
Modified-site	/label= putative N-glycosylation site
Modified-site	802..804
Modified-site	/label= putative N-glycosylation site
Modified-site	881..883
Modified-site	/label= putative N-glycosylation site
Modified-site	901..903
Modified-site	/label= putative N-glycosylation site
Modified-site	912..914
Modified-site	/label= putative N-glycosylation site
Modified-site	941..943
Modified-site	/label= putative N-glycosylation site
Modified-site	947..949
Modified-site	/label= putative N-glycosylation site
Modified-site	979..981
Modified-site	/label= putative N-glycosylation site

FT Modified-site 994..996
 FT /label= putative N-glycosylation site
 FT 1022..1024
 FT Modified-site
 FT /label= putative N-glycosylation site
 FT 1045..1047
 FT Modified-site
 FT /label= putative N-glycosylation site
 FT 1051..1053
 FT Modified-site
 FT /label= putative N-glycosylation site
 FT 1076..1078
 FT Modified-site
 FT /label= putative N-glycosylation site
 FT 1106..1134
 FT /label= putative_transmembrane_region
 XX
 PN EP364690-A.
 XX
 PD 25-APR-1990.
 XX
 PF 17-AUG-1989; 89EP-00115159.
 XX
 PR 23-AUG-1988; 88US-00235353.
 PR 09-MAR-1989; 89US-00321239.
 XX
 PA (DAND) DANA FARBER CANCER INST INC.
 XX
 PI Springer TA, Corbi A;
 XX
 DR WPI: 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 PT New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
 PT inflammation and viral infections, and in diagnosis.
 PT
 PS Disclosure; Page ?; 3pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
 CC -MAR-2003 to correct PA field.)
 CC
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key
 CC and pages
 CC
 XX Sequence 1153 AA;
 SQ
 Query Match 99.1%; Score 5829; DB 2; Length 1153;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1123; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 60
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 76
 QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSLNRQQPQK 120
 DB 77 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSLNRQQPQK 136
 QY 121 PPEALRGCPQESDIAFLIDGSGSIIPDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
 DB 137 PPEALRGCPQESDIAFLIDGSGSIIPDFRMKELVSTIMEQLKSKTFLSLMOYSEEF 196
 QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRTHTTATGKRVVRELFNITNGARKNAKILFL 240
 DB 197 RIHFTFKFQNNPNSRLIKPTQLLGRTHTTATGKRVVRELFNITNGARKNAKILVVI 256
 QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN 300
 DB 257 TDGEKFGDPLGYEDVPEADRGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN 316
 QY 301 NFEALKTVONQIREKIFAIEGTQTGSSSFEHMQSGEFSAAITNGPLLSTVGYDWDAG 360
 DB 317 NFEALKTVONQIREKIFAIEGTQTGSSSFEHMQSGEFSAAITNGPLLSTVGYDWDAG 376

QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYOHIGLVAMFR 420
 DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYOHIGLVAMFR 436
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 496
 QY 481 PRGQARWQCDVILYGEQGPWGFRGAAITVLGVNGDKLTDVAIGAPGEEDNGAVYLF 540
 DB 497 PRGQARWQCDVILYGEQGPWGFRGAAITVLGVNGDKLTDVAIGAPGEEDNGAVYLF 556
 QY 541 HGTSGSISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGTSGSISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
 DB 617 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSRTOVGLGTOTCETLKLQLPNCIEDPVSPIVRLNF 720
 DB 677 YDLALDSGRPHSRVAFNETKNSRTOVGLGTOTCETLKLQLPNCIEDPVSPIVRLNF 736
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNGNDNIQDDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNGNDNIQDDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
 DB 797 GPREFNVTTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQSRSLPISLVFLVPV 960
 DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQSRSLPISLVFLVPV 976
 QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVVVNCISIAVCQRIQCDIP 1020
 DB 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVVVNCISIAVCQRIQCDIP 1036
 QY 1021 PFGIOEEFNATLKGNLSEFDWIKTSHNHLIVSTAEILFNDVSFTLLPQGGAFVRSQDET 1080
 DB 1037 PFGIOEEFNATLKGNLSEFDWIKTSHNHLIVSTAEILFNDVSFTLLPQGGAFVRSQDET 1096
 QY 1081 KVEPPEVENPLPLIVGSSVGGLLILALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
 DB 1097 KVEPPEVENPLPLIVGSSVGGLLILALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153
 RESULT 8
 ID ADM99589 standard; protein; 1152 AA.
 XX
 AC ADM99589;
 XX
 DT 17-JUN-2004 (first entry)
 DE Human integrin alphaM subunit precursor protein.
 XX
 KW integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
 KW osteopathic; cytostatic; immunosuppressive; anti-inflammatory;
 KW neuroprotective; antiskinning; immunotherapy; cancer; osteoporosis;
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
 KW alphaM.
 OS Homo sapiens.
 XX

PH	Key	Location/Qualifiers	
FT	Misc-difference 965	/note= "Encoded by CCC"	
FT			
XX			
PN	WO2004007530-A2.		
XX			
PD	22-JAN-2004.		
XX			
XX	17-JUL-2003; 2003WO-US022301.		
XX			
PR	17-JUL-2002; 2002US-0396783P.		
PR	17-JUL-2002; 2002US-0396790P.		
PR	11-SEP-2002; 2002US-0410135P.		
XX			
XX	(BLOO-) CENT BLOOD RES INC.		
PA			
XX			
PI	Springer TA, Takagi J;		
XX			
XX	WPI; 2004-122877/12.		
DR	N-PSDB; ADM99588.		
XX			
PT	Novel modified integrin protein having extracellular domains of integrin alpha and beta subunits or integrin alphaII and beta3 subunit, useful for treating integrin mediated disorders.		
PT			
PT			
XX			
XX	Disclosure; SEQ ID NO 4; 232pp; English.		
XX			
CC	The invention relates to a novel isolated or recombinant modified integrin protein having extracellular domains of integrin alpha and beta subunits where one of the subunits has one or more mutations, an altered surface feature or an amino acid substitution or internal deletion, extracellular domains of the integrin beta subunit that comprise a mutation that alters a non-cysteine residue to cysteine or extracellular domains of integrin alpha and beta subunits. The polypeptide of the invention demonstrates antiproliferative, thrombolytic, anticoagulant, osteoprotective, cyostatic, immunosuppressive, antiinflammatory, neuroprotective and antiskinning activities and may be useful for immunotherapy in order to prevent or treat an integrin-mediated disorder such as an inflammatory disorder, an autoimmune disorder, thrombosis, cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple sclerosis. The current sequence is that of the human integrin alphaM subunit precursor protein of the invention.		
XX			
XX	Sequence 1152 AA;		
	Query Match	99.1%; Score 5823.5; DB 8; Length 1152;	
	Best Local Similarity	98.8%; Pred. No. 0;	
	Matches 1123; Conservative	8; Mismatches	5; Indels
			1; Gaps
QY	1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI	60	
DB	17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI	76	
QY	61 RLQVPVAVNMVSLGLSLAATTPSPQLACGPTVHQTCSNTYKGLCFLFGSNLRQOPQK	120	
DB	77 RLQVPVAVNMVSLGLSLAATTPSPQLACGPTVHQTCSNTYKGLCFLFGSNLRQOPQK	136	
QY	121 FPEALRCQPEDSDIAPLIDGSGIIPHDPRRMKELVSTIMEQLKSKTFLPSLMQYSEEF	180	
DB	137 FPEALRCQPEDSDIAPLIDGSGIIPHDPRRMKELVSTIMEQLKSKTFLPSLMQYSEEF	196	
QY	181 RIHPTFKFQNNPRSLRIPITOLLGRTHATGLRKVRVRELFINITNGARKNAFKILFL	240	
DB	197 RIHPTFKFQNNPRSLRIPITOLLGRTHATGLRKVRVRELFINITNGARKNAFKILV	256	
QY	241 TDGEKFGDPLGYEDVPELDRGVIRYVVGDAFRSEKSRQELNIVASKPPRDHVFQAN	300	
DB	257 TDGEKFGDPLGYEDVPELDRGVIRYVVGDAFRSEKSRQELNIVASKPPRDHVFQAN	316	
QY	301 NFEALKTVQNLREKIPFAIGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	360	
DB	317 NFEALKTVQNLREKIPFAIGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	376	

QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVGAPRYOHIGLAVMER	420
DB	377	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVGAPRYOHIGLAVMER	436
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL	480
DB	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL	496
QY	481	PRGORARQCDAVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEEDNRGAVYLF	540
DB	497	PRG-RARQCDAVLYGEGQGPWGRFGAALTVDVGVNGDKLTDVAIGAPGEEDNRGAVYLF	555
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTWDGLVDLTVGAGHVLLRSQ	600
DB	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTWDGLVDLTVGAGHVLLRSQ	615
QY	601	PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGIOQSVVT	660
DB	616	PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGIOQSVVT	675
QY	661	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF	720
DB	676	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRNF	735
QY	721	SLVGTPLSAFGLRPLVLAEDAQRLEFALPFPEKXCGNDNICODDLSTIFFSMSLDCLVVG	780
DB	736	SLVGTPLSAFGLRPLVLAEDAQRLEFALPFPEKXCGNDNICODDLSTIFFSMSLDCLVVG	795
QY	781	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWRLACESASSTEV	840
DB	796	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWRLACESASSTEV	855
QY	841	SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
DB	856	SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	915
QY	901	QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMOHQYQVSNLGRSLPISLVLVLPV	960
DB	916	QLELPVKYAVYVMTSHGVSTKYLNTASENTSRVMOHQYQVSNLGRSLPISLVLVLPV	975
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCORIQCDIP	1020
DB	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCORIQCDIP	1035
QY	1021	FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET	1080
DB	1036	FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAEILFNDSVFTLLPGQAFVRSQTET	1095
QY	1081	KVEPFEVNPPLIVGSSVGLLLALLITAAALYKLGFFKQYKDMKSEGGPPGAEPQ	1137
DB	1096	KVEPFEVNPPLIVGSSVGLLLALLITAAALYKLGFFKQYKDMKSEGGPPGAEPQ	1152
	RESULT 9		
	ADP12435		
ID	ADP12435	standard; protein; 1152 AA.	
XX	ADP12435;		
AC	ADP12435;		
XX	ADP12435;		
DT	12-AUG-2004 (first entry)		
XX			
DE	Protein encoded by mRNA of the invention #45.		
XX			
KW	transplant rejection; immune system; rheumatoid arthritis; lupus;		
XX	inflammatory bowel disease; multiple sclerosis; HIV; AIDS.		
OS	Homo sapiens.		
XX			
PN	WO2004042346-A2.		
PD	21-MAY-2004.		
XX			
PF	24-APR-2003; 2003WO-US012946.		

XX 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
PA Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX WPI; 2004-400724/37.
DR Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2444; 1762pp; English.
XX The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The methods are also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 1152 AA;
Query Match 98.9%; Score 5816.5; DB 8; Length 1152;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 8; Mismatches 6; Indels 1; Gaps 1;
QY 1 ENLDTENAMTFOENARGGQSVVQLQGGVVVVGAPQEIIVAAANQGSLSYQCDYSTGSCPEI 60
DB 17 ENLDTENAMTFOENARGGQSVVQLQGGVVVVGAPQEIIVAAANQGSLSYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQQPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQQPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLPSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLPSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVRELFTNITNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVRELFTNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDGAFRSEKSRQELNTVASKPPRDRHVQAN 300
DB 257 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDGAFRSEKSRQELNTVASKPPRDRHVQAN 316
QY 301 NPEALKTVQNLQREKIFAIEGTQGTSSSSFEHMSQEGHSAITNSGPLLSVTGSDWAG 360
DB 317 NPEALKTVQNLQREKIFAIEGTQGTSSSSFEHMSQEGHSAITNSGPLLSVTGSDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILRNVRQSLVLGAPYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILRNVRQSLVLGAPYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQSVCP 496
QY 481 PRGQARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 497 PRGQARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 555

QY 541 HGTSGSGTSPSHSRIAGSKLSRLQYFGQSLSGQDLTMDGLVDLTVGAQHVLILRSQ 600
DB 556 HGTSGSGTSPSHSRIAGSKLSRLQYFGQSLSGQDLTMDGLVDLTVGAQHVLILRSQ 615
QY 601 FVLRVKAIIMEFNPREVARNFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 616 FVLRVKAIIMEFNPREVARNFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPSRAVFNETHKSTRQTOVLGTCETLKLQPLNCIEDPVSPIVRLNF 720
DB 676 YDLALDSGRPSRAVFNETHKSTRQTOVLGTCETLKLQPLNCIEDPVSPIVRLNF 735
QY 721 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPPEKKGNDNIQDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFAGNLRLPVLAEADQRLFTALFPPEKKGNDNIQDDLSITFSFMSLDCLVVG 795
QY 781 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 796 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 960
DB 916 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020
DB 976 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1035
QY 1021 PFGIOEEENATLKGNSLFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQDET 1080
DB 1036 PFGIOEEENATLKGNSLFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQDET 1095
QY 1081 KVEPPEVENPFLIVGSSVGGILLIALITALYKLIGFFKQYKDMWSEGGPPGABPQ 1137
DB 1096 KVEPPEVENPFLIVGSSVGGILLIALITALYKLIGFFKQYKDMWSEGGPPGABPQ 1152
RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX ADP44061;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human CD11C protein SEQ ID NO:14.
XX
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX Homo sapiens.
XX WO2004053094-A2.
XX
PD 24-JUN-2004.
XX
PF 08-DEC-2003; 2003WO-US039208.
XX
PR 06-DEC-2002; 2002US-0431522P.
XX (PPDP-) PPD DEV LP.
PA Dunn SJ, Holzwayer TA;
XX PI WPI; 2004-480928/45.
XX N-PSDB; ADP44060.
DR
XX Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

PT agent for IBD.
XX Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC The invention comprises marker genes for inflammatory bowel disease (IBD)
CC - FcgammaRIIa, FcgammaRIIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC The DNA and protein sequences of the invention are useful in the
CC detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC disease). The present amino acid sequence represents an IBD marker
CC protein of the invention.
XX
XX Sequence 1163 AA;
Query Match 58.6%; Score 3446; DB 8; Length 1163;
Best Local Similarity 60.5%; Pred. No. 5.4e-276;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;
1 FNLDENAMTFOENARGFGQSVOLQGSRRVVVGAPQETVAANQSGSLYQCDYSTGSCPEPI 60
20 FNLDTEELTAFRVDSAGGDSVVQYANSVVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
61 RLQVPVEAVNMSLGLSLAATTSPQLLACGTVHTQTCENTYVKGLCFLFGSNLRQQQX 120
80 GLQVPPEAVNMSLGLSLAATTSPQLLACGTVHTQTCENTYVKGLCFLFGSNLRQQQX 137
121 FPEALRGCPQSDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
138 LPVSRQECPRQEDIVFLIDGSGISSRNFATMNFVRAVISQFQPSQTSMLQFQSNKF 197
181 RIHTTFKFFONPNRSLIKIPITOLLGRTHATGLRKVVRELFNITNGARKNAKILPLL 240
198 QTHTFEERFTSNPLSLASVHQLQGFYTAIQLNVVHRLFNHSGYARDAIKILIVI 257
241 TDGKFGDPLGYEIVPELDREGVIRYVGLGDPAPRSEKSKQELNVTASKPRDHVFOAN 300
258 TDGKKGSDLDYKDVIPMADAAGIIRYAGVGLAFQNRNSWKNELNDIASKQSEHIFKVE 317
301 NFEALKTVQNLREKIFAIEGTQTCSSSEFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
318 DFDALKIOQLREKIFAIEGTQTCSSSEFEHMSQEGFSAITNSGPLLSTVGSYDWAG 377
361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAILNRNQSLLVGLGAPYOHIGLVAMFR 420
378 GAFLYPPNMSPTFINMSQENVMDRDSYLGYSYTELALWKGVQSLVGLGAPYOHIGLVAMFR 437
421 QNTGWSNANVKTQIGAYFGASLCSDVDNSGSDTLVLICAPHYEQTRGGQVSVCP 480
438 QVSRQWRMKAETVGTQIGSYFGASLCSDVDNSGSDTLVLICAPHYEQTRGGQVSVCP 497
481 PRGQARWCDAVLYGEGQGPWRFGAALTVDGVNGDKLTDVATGAPGEEDNRGAVYLF 540
498 PRGWR-RWCCDAVLYGEGQGPWRFGAALTVDGVNGDKLTDVATGAPGEEDNRGAVYLF 556
541 HGTSGSISPSHSORIASGKSLPRIQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 600
557 HGVLGPSISPSHSORIASGKSLPRIQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 616
601 PVLRYKALMENPREVARNVECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLWVGSMQIPAIPEIPSAFECREQVYSEQTLVQSNICLYIDKRNKLLSGRDLQSSVT 676
661 YDLALDSRPHSRVFNETKNSRTOVLGLTQTCETLKLQPLNCIEDPVSPIVLRNF 720
677 LDALDLPORLSPRATFQTKRSLSRVVLGLKAHCENFNLLPSCVSDSVTPITLRNF 736
721 SLVGTPLAFAGNLRLVLAEDARLFTALFPPEKNGNDNIQDDLSITTFMSDCLVVG 780
737 TLVGPFLAFAFNRLRMLAALQRYFTASLPPEKNGADHIQDNLGIGSFPGKSLVVG 796
781 GPREFNVTVTVDNEDSDSYRTQVTFPPFLDLSYKRVSTLQNRORSWRLACESASSTEV 840
797 SNLELNAEYVMWVNDGEDSYGTTITTFSHPAGLSYRYVAEGQKQQLRSLHLTCDSPAVG-- 854

QY 841 SGALKSTSCSINHPIPPENSEVTENITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
DB 855 SQGTWSTSCRINHPIFRGGAQITFLATFDVSPKAVLGRLLLTANVTSENNMPTSKTTF 914
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFTAS-ENTSRVMQHOYQVSNLQORSILFISLVFLVP 959
DB 915 QLELPVKYAVVMVVTSHGVSTKYLNFTAS-ENTSRVMQHOYQVSNLQORSILFISLVFLVP 974
QY 960 VRLNQTVMWRDPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNCISIAVCORIQCDI 1019
DB 975 VELNQAAMVMDVEVSHPNPSLRCSSEKIAFPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034
QY 1020 PFGIOEFENATLKNLSFDWYIKTSHNHLIVSTAEILNDSVFTLLPGGAFVRSOTE 1079
DB 1035 PSFSVGELEDFTLKNLSFGWVRQLOKKVSVSVVAEITFDTSVYSQLPGEAFMRAGIT 1094
QY 1080 TKVPEPFEVNPPLPLIVGSSVGLLLALITAALYKLGFFKQYKDMSE 1128
DB 1095 TVLEKYKVHNPPLPLIVGSSVGLLLALITAALYKLGFFKQYKDMSE 1143
RESULT 12
ADQ17510
ID ADQ17510 standard; protein; 1163 AA.
XX ADQ17510;
AC ADQ17510;
DT 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
KW Homo sapiens.
OS WO2004048938-A2.
PN 10-JUN-2004.
XX 26-NOV-2003; 2003WO-US038193.
XX 26-NOV-2002; 2002US-0429739P.
XX (PROT-) PROTEIN DESIGN LABS INC.
PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPT; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX Example 2; SEQ ID NO 327; 210pp; English.
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cycostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX Sequence 1163 AA;
SQ

Query Match		58.6%; Score 3446; DB 8; Length 1163;
Best Local Similarity		60.5%; Pred. No. 5.4e-276;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;		
QY	1	FNLTEAMTQENARGGQSVVQLQSGRRVVGAPQEIIVAAQNRGLYQCQDYSTGSCPEI 60
Db	20	FNLTEELTAFRVDSAGSGDSVVQYANGSVVVGAPKITAANQTGGLYQCGYSTGACEPI 79
QY	61	RLOVPVAVNMNLGLSLAATTSPOLLACGTVHQTCSNTYVVGKLCFLFGSNLRQDPQK 120
Db	80	GLQVPEAVNMNLGLSLASTTSPOLLACGTVHHECGRNMYLTGLCLFLLGPT--QLTR 137
QY	121	FPEALRGCPQSDSIAFLIDSGSIIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db	138	LPVSRQECPRQEQDIVFLIDSGSISSENFMNFRVRAVISQFQRPSTQSLMQFSNKF 197
QY	181	RIHTFKFQNNPNRSLIKPITOLLGRHTATGLRKVKVRELFMNITNGARKNAFKILFLL 240
Db	198	QTHFTFEFRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257
QY	241	TGGEKFGDPLGYEDVPELDEGVIRYVVGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db	258	TGKKEGDGLDYKDVIMADAGIIRYAIGVGLAFQNRNSWKELNDTASKEQBHFKEVE 317
QY	301	NFEALKTQVQNLREKIFAIEGTQTCSSSSFFHEMSQEGFSAAITSNGLLSTVGSYDWAG 360
Db	318	DFDALKDIQNLKEKIFAIEGTEITSSSSFFLEWAQEGFSAVFTPDGVLGAVGSFTWSG 377
QY	361	GVFLYTSKEKSTFNMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db	378	GAFLYPPNMSPTFINMSQENVMRDSVLGYSTELALWKGVQSLVGLGAPRYOHTKAVIFT 437
QY	421	QNTGQWESNANVKCTQTCGAVGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQVSCPL 480
Db	438	QVSRQWMAKAEVTCQIGSVFASLCSVDVDTGSDTLVLIGAPHYVEQTRGGQVSCPL 497
QY	481	PRGQARWQCDVAVLYGQGPWGRFGAALTVLGVDWNGDKLTDVAIGAPGEDNAGAVLYF 540
Db	498	PRGWR-RWDCDVLVYGGQHPWGRFGAALTVLGVDWNGDKLTDVVGAPGEENRAGAVLYF 556
QY	541	HGTSGSISPSHSQRIAGSKLSPLOVFGQSLSGQDLTMDGLVDLTVGAGQHVLLRSQ 600
Db	557	HGVLGPISPSHSQRIAGSQLSSRLQVFGQALSGQDLTQDGLVDLAVGARGQVLLLR 616
QY	601	PVLVRKAIMENPREVARNVPECNDVYVKGAEVRCVCHLVQKSTDRLRREGIOISVVT 660
Db	617	PVLVWGVSMQFIPIAIFRPAFECEQVVBQTLVQSNICLYIDRKNLLGSRDLQSVT 676
QY	661	YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETKLQLPNCIEDPVSPTVLRINF 720
Db	677	LDLALDPGLSPRATFOETKNRSLSRVRLGLKAHCENFLLLPSCVEDSVTPITLRINF 736
QY	721	SLVGTPLSAFGRNRPVLAEDAQRIFTALPFPEKNCNDNIQQDLSITFSFMSLDCIVVG 780
Db	737	TLVGKPLLAFLNRLPMLAALQRYFTASLPFEKNCAGADHICQDNLGIFSFPGLKSLVG 796
QY	781	GPREFNVTVVRNDGEDSVYQTVTFEPDLDSYRKVSTLQNRQSQRWRLACBSASSTEV 840
Db	797	SNLELNEVWMDGEDSYGTTITFSHPAGLSYRYVAEGQKQQLRSLHLTCDSPVYG-- 854
QY	841	SGALKSTSCSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db	855	SGQWTSCTRINHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSENNTPRTSKTTP 914
QY	901	QLELPVKYAVYVTSHGVTSKYLNFTAS-ENTSRVMOHQYQVSNLQGRSLPISLVLVP 959
Db	915	QLELPVKYAVYTVVSSHEQTKYLNFSSEKESHVAMHYQVNNLQORDLPVSINFWVP 974
QY	960	VRNLQTVIMDRPQVTFSENSSICHTXERLPSHSDFLAELRKAPVNCSTAVCQRCQDI 1019
Db	975	VELNQEAVMWDMVESHQPNFLCSSEKIAAPPASDFLAHIQKNPFLVDCSIAGCLRFCDV 1034
QY	1020	PFFGIQEFENATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGQGAFFVRSQTE 1079
Db		

Db	1035	PSFSVQBELDTLKGNSLFGWVRQLQKKVSVVAETFTDTSVYSQLPGQEAFFRAQTT 1094
QY	1080	TKVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSE 1128
Db	1095	TVLEKYKVHNPPLIVGSSIGLLLLALITAVLYKVGFFKQYKEMEE 1143
RESULT 13		
AAAR07120	ID AAR07120 standard; protein; 1163 AA.	
XX	AC	AAAR07120;
XX	DT	25-MAR-2003 (revised)
XX	DT	05-FEB-1991 (first entry)
XX	DE	p150.95 alpha subunit encoded by clone lambdaX47.
XX	KW	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia; rhinovirus.
XX	OS	Synthetic.
XX	XX	Key Location/Qualifiers
FT	Region	1..19
FT	Region	/label= signal peptide
FT	Region	20..44
FT	Region	/label= N-terminus
FT	Modified-site	61..63
FT	Modified-site	/label= glycosylation site
FT	Modified-site	89..91
FT	Modified-site	/label= glycosylation site
FT	Modified-site	385..387
FT	Modified-site	/label= glycosylation site
FT	Modified-site	392..394
FT	Modified-site	/label= glycosylation site
FT	Modified-site	697..699
FT	Modified-site	/label= glycosylation site
FT	Modified-site	735..737
FT	Modified-site	/label= glycosylation site
FT	Modified-site	899..901
FT	Modified-site	/label= glycosylation site
FT	Modified-site	904..906
FT	Modified-site	/label= glycosylation site
FT	Modified-site	939..941
FT	Modified-site	/label= glycosylation site
FT	Modified-site	1050..1052
FT	Modified-site	/label= glycosylation site
FT	Domain	1108..1133
FT	Domain	/label= transmembrane
XX	XX	WO9010646-A.
XX	XX	PN
XX	XX	20-SEP-1990.
XX	XX	09-MAR-1990; 90WO-US001257.
XX	XX	09-MAR-1990; 90WO-US001257.
XX	XX	(DAND) DANA FARBER CANCER INST INC.
XX	XX	Corbi AA, Springer TA;
XX	XX	WPI; 1990-304985/40.
XX	XX	N-PSDB; AAQ06068.
XX	XX	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-unit of p150.95 cell surface adhesion receptor, opt. together with a beta chain of CD-18 family.
XX	XX	Disclosure; Fig 3; 59pp; English.

CC Clone lambda X47 was isolated from a cDNA library constructed from total
CC RNA extracted from phorbol myristate acetate stimulated HL-60
CC myelomonocytic cells. The library was screened with oligonucleotide
CC probes based on tryptic peptide fragments of p150.95. The sequence can be
CC attached to appropriate control elements and expressed in prokaryotic and
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC infection because it interacts with ICAM-1 and inhibits cell-virus
CC attachment. It can also be used as an anti-inflammatory agent. See also
CC AAQ06063-4, AAQ06069, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
xx
xx Sequence 1163 AA;

QY	841	SGALKSTCSINHP	IPENSEVTFNTITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
Db	855	SQTWSTSCRINHL	IFRFGAQITFLAFDVSPKAVIGDRLLLTANVSSENTPRTSKTTF	914
QY	901	QLELPVKYAVY	VMVTVTSHGYSTKYLNFTAS-ENTSRVMQHOYQVSNLQGRSLPISLVLFP	959
Db	915	QLELPVKYAVY	TVVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQGRDLPVSNFWVP	974
QY	960	VRNLQTVINDR	POVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDI	1019
Db	975	VELNLQAVMMD	VEVSHPOPNPSLRCSQKIAPPASDDELAHQKNPVLDCSIAGCLRFCDV	1034
QY	1020	PFTGIOBEFNATL	KGNLSFDWYIKTSHNHLIIYSTAEILFENDSVFTLLPQCGAFVRSQTE	1079
Db	1035	PSFSVQELDF	TLKGNLSFGWVQILQKKVSVVVAEITDTSVYSQLPQGERFMAQTT	1094
QY	1080	TKPEPPEVNP	PLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMNSE	1128
Db	1095	TVLEKYKVHNPT	PLIVGSSIGALLLALITAVILYKVGFFKRYKEMMEE	1143
RESULT 14				
AAW65091	ID	AAW65091	standard; protein; 1163 AA.	
XX	AC	AC	AAW65091;	
XX	DT	28-SEP-1998	(first entry)	
XX	DE		Human Beta-integrin CD11c subunit protein.	
XX	KW		Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;	
XX	KW		type-I diabetes; atherosclerosis; multiple sclerosis; asthma;	
XX	KW		lung inflammation; acute respiratory distress syndrome; CD11c subunit;	
XX	KW		rheumatoid arthritis.	
XX	OS		Homo sapiens.	
XX	XX		US5728533-A.	
XX	PN		17-MAR-1998.	
XX	PD			
XX	PF		07-JUN-1995; 95US-00485618.	
XX	PR		23-DEC-1993; 93US-00173497.	
XX	PR		05-AUG-1994; 94US-00286989.	
XX	PR		21-DEC-1994; 94US-00362652.	
XX	XX		(ICOS-) ICOS CORP.	
XX	FA			
XX	PI		Van Der Vieren M, Gallatin WM;	
XX	XX		WPI; 1998-206565/18.	
XX	DR			
XX	XX			
PT	PT		Screening assay for modulators of integrin binding - using immobilised or	
PT	PT		labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.	
XX	XX			
PS	PS		Example 5; Fig 1A-D; 106pp; English.	
CC	CC		This sequence represents a human beta-integrin CD11c subunit which is	
CC	CC		used to describe a method for identifying compounds that modulate the	
CC	CC		interaction of the beta-integrin alpha-d subunit with a binding partner	
CC	CC		of alpha-d which involves contacting an alpha-d polypeptide with an alpha	
CC	CC		-d binding partner, one of which is immobilised and the other of which is	
CC	CC		labelled, in the presence of a test compound, and determining if the	
CC	CC		compound affects binding between the alpha-d polypeptide and alpha-d	
CC	CC		binding partner, where the alpha-d polypeptide is alpha-d or its fragment	
CC	CC		comprising the cytoplasmic, transmembrane or extracellular domain of	
CC	CC		alpha-d. Compounds that modulate alpha-d binding could be used to treat	
CC	CC		diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,	
CC	CC		asthma, psoriasis, lung inflammation, acute respiratory distress syndrome	
XX	XX		and rheumatoid arthritis	

SQ Sequence 1163 AA;
 Query Match 58.2%; Score 3423; DB 2; Length 1163;
 Best Local Similarity 60.2%; Pred. No. 4.4e-274;
 Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4

QY	1	ENLDTENAMTQENARGFGQSVVQLGGRRVVVAGPQEIIVAAQNQRSLYQCYDYSTGSCBPI	60
DB	20	FNLDEELTAFRVDSAGFGSDVQVANSVVVWVGAPOKIIAANQIGGLYQCYGYSTGACBPI	79
QY	61	RLQVPEAVANMGLSLAAATTPSPOLLACGPTVHCTCSENTYVKGILCLFGSLNLRQOQOK	120
DB	80	GLQVFPFPAVNMSLGLSLASTTSPSLQACGPTVHHCGRNMYLTGLCLLLGPT--QLTQR	137
QY	121	FPEALRCPOEDSDIAFLIDGSGSTIHPDFFRMKELVSTIMEQLKKSTLFSLMOYSEEF	180
DB	138	LPVSRQCPREQDQIVFLIDGSGSISSRFATMMNFRAVISOQRPSTQSLMQFSKNF	197
QY	181	RIIHFTPEFQNNPNSRLIKPITOLLGRTHTATGLRKVVRELFNITNGARKNAFKILFL	240
DB	198	QTHFTPEEFRTSNPLSILASVHQLQGTPTYATAQNVVHRLFHASYGARRDAIKILVI	257
QY	241	TDEKFGDPLGYEDVIPLEDEGVIRYVLGFGDAFRSEKSRQELINTVASKPRDHVFOAN	300
DB	258	TDGKKGSDLDYKDVIPMAADAGIIRVAIGVLAFQPNRNSWKEKLNIDASKESQSHIFKE	317
QY	301	NFEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG	360
DB	318	DFDALKDIQNLKEKIFAIEGTETIISSSFLEWAQEGFSVTFDGPVLGAVGSGFTWSG	377
QY	361	GVFLYTSKEKSTFNNTRVDSDDNDAYLGAAAAIILNRVOSLVLGAPRYQHIGLVAMFR	420
DB	378	GAFLYPENMGPTTFINMGQENVDMDSYLGSTELALWKGVSQVLGAPRYQHIGKAVIFI	437
QY	421	QNTGMWESNANVKCTQLGAYFGASLCSVDVDSNGSTDLVLGAPHYYEOTFGGVSVCP	480
DB	438	QVSRQWPKAEVIGTQLGSTFGASLCSVDVDTGSTDVLVGAPHYYEOTFGGVSVCP	497
QY	481	PRGQARWQCDADVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF	540
DB	498	PRGWR-RWMCADVLYGEGHPWGRFGAALTVLGVDVNGDKLTDVVI GAPGEENRGAVYLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLYFGQSLGGQDLTMDGLVDLTVGAGHVLRLRSQ	600
DB	557	HGVLGPIPSPSHSQRIAGSLSRQLYFGQALSGQDLTDGLVDLAVGARGQVLLLR	616
QY	601	PVLVRKAIMEFNPREVARNVFECDNDVVYKGEAGEVRVCLHVQKSTDRDLRBEQIQSVVT	660
DB	617	PVLVVGVSQFI PAEIPRSAFECEQVVSQTLVQSNICLYIDKRKSNLLGSLRDLQSSVT	676
QY	661	YDIALDSGRPHSRVAFNETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPTVLRLNF	720
DB	677	LDLALAPGRLSPRAIQOETKNRSLSRVVLGLKAHCENFNLLPSCVEDSVIPIILRLNF	736
QY	721	SLVGTPLSAPGNLRPVLAEDAQRLEFTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVVG	780
DB	737	TLVGKPLLAFLNLRPMLAALQAQRYFTASLPFEKNCAGADHICQDNLGISFSFPFGLKSLVG	796
QY	781	GREPEFNVTVTRNDGEDSYRTQTFPPPLDLVRKVVSTLQNRQSRWRRLACESASTEV	840
DB	797	SNUELNAEVMVWMDGEDSYGTTITFHPAGLSRYVAEGQKQGRJSLHTC--CSAPVG	854
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENKMPRTNKTFF	900
DB	855	SQCTWSTSCINHLIFPGGAQIITFLATFDVSPKAVGLDRLILLIANVSENNIIRTSKTI	914
QY	901	QLELPVKYAYVMVTVSHGVSTKYLNFPTAS-ENTSRVMHQYQVSNLQGRSLIPSLVFLVP	959
DB	915	QLELPVKYAYIVVSSHEQTKYLNFSSEKESHVAMHRYQVNNLQORDLPVSIINFPVP	974
QY	960	VLRLNQIVIMDRPOVITSENLSSTCHTKERIPLSHSDFLAELRKAPVNVNCSIAVCQRIQCDI	1019
DB	975	VELNQBAVMMDVRSHPQNPDLRSCSEKIPAPSDFLAHTQKQNPVLDCSLTAGCIRFCDV	1034

QY	1020	PFFGIQEBFNATLKGNSLFDWIKTSNNHLLIIVSTAELFNDVSFTLLPGQAFVRSQTE	1079
DB	1035	PSFSVQOEELDTFLKGNLSFGWVRQILQKSVSVSVAEIIIDTSVYSQLPGQAFMRAQTI	1094
QY	1080	TKVZPFEVNPPLPIVGVSSVGGLLLLALITAAALYKGLGFFKROYKDMWSE	1128
DB	1095	TVLEKYVHNPIPLVGVSSVGGLLLLALITAVLYKVGFFKROYKEMWEE	1143
RESULT 15			
ID	AAB07361		
XX	AAB07361 standard; protein; 1163 AA.		
AC	AAB07361;		
DT	17-JAN-2001 (first entry)		
DE	Human CD11c protein sequence.		
KW	Human; macrophage infiltration inhibition; alpha_d integrin;		
KW	leukocyte integrin; Leu-CAM; leukointegrin; immune response;		
KW	inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;		
KW	atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;		
KW	lung inflammation; acute respiratory distress syndrome; Crohn's disease;		
KW	rheumatoid arthritis; central nervous system injury; CD11c.		
OS	Homo sapiens.		
XX	WO200029446-A1.		
XX	25-MAY-2000.		
XX	16-NOV-1999; 99WO-US027139.		
XX	16-NOV-1998; 98US-00193043.		
PR	08-JUL-1999; 99US-00350259.		
PA	(ICOS-) ICOS CORP.		
PI	Gallatin MW, Van Der Vieren M;		
XX	WPI; 2000-387751/33.		
PT	Use of novel anti-alpha integrin d monoclonal antibodies to inhibit		
PT	macrophage infiltration and reduce inflammation at central nervous system		
XX	injury sites.		
PS	Example 5; Fig 1; 270pp; English.		
CC	Integrins are a class of membrane-associated molecules that participate		
CC	in cellular adhesion. Integrins are made up of an alpha subunit and a		
CC	beta subunit. One class of human integrins are restricted to expression		
CC	in white blood cells and have a common beta2 subunit: the leukocyte		
CC	integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins		
CC	have an important role in immune and inflammatory responses. The present		
CC	protein sequence is the human integrin alpha subunit CD11c. This sequence		
CC	was used in an alignment to identify a novel beta2 integrin alpha		
CC	subunit: alpha d (AAA60014 and AAB07359). The present sequence has		
CC	approximately 66% identity to the protein sequence of alpha d. The		
CC	Alpha d gene and protein may be useful in therapy for diseases linked to		
CC	aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple		
CC	sclerosis, asthma, psoriasis, lung inflammation, acute respiratory		
CC	distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency		
CC	(LAD). In addition, anti-alpha d monoclonal antibodies may be used in the		
CC	inhibition of macrophage infiltration at the site of a central nervous		
CC	system injury. The monoclonal antibodies can also be used to detect and		
CC	diagnose Crohn's disease		
XX	Sequence 1163 AA;		
XX	Sequence 1163 AA;		

Query Match 58.2%; Score 3423; DB 3; Length 1163;
 Best Local Similarity 60.2%; Pred. No. 4.4e-274;

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;			
Qy	1	FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNGSLVQCDYSTGSCPEPI 60	
Db	20	FNLDEELTAFRVDSAGFGSDVQVANSVWVGAPQKIIAANQIGGLYQCGYSTGACEPI 79	
Qy	61	PLQVPEAVNWSLGLSLAATSPPOLACGPTVHQTCSNTYVKGCLFLGSLNLRQOPQK 120	
Db	80	GLQVPEAVNWSLGLSLASTTSPQLACGPTVHHCGRNMYLTGLCLFLGPT--QLTQR 137	
Qy	121	FPREALRGCPQSDSDIAFLIDSGSIIIPHDFFRMKELVSTIMEQLKKSKTLFSLMQYSEEP 180	
Db	138	LPVSRQECPRQEQDIVFLIDSGSISRNFAFMNFVRAVISQFORPSTQFSLMQFSNKF 197	
Qy	181	RIHFTFKFQNNPNRPSIKPITQLLGRTHATGLRKVVURELFNITNGARKNAFKILFLL 240	
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGFTYTATAIQNVVHRLPHASYGARRDAIKILIVI 257	
Qy	241	TDGEKFGDPLGYEDVPELDRGVIRYVLFGDFAFRSEKSRQELNTVASKPPDRDHVFOAN 300	
Db	258	TDGKKEGSDLDYKQVIMADAAGIIRYAIQVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317	
Qy	301	NPEALKTVQNLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360	
Db	318	DFDALKDQNLKQKIFAIEGTETITSSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG 377	
Qy	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNVRQSLVLGAPRYOHIGLVAMPR 420	
Db	378	GAFLYPPNMSPTFINNSQENVDMRDSYLGSTELALWKGVQSLVLGAPRYOHIGKAVIFI 437	
Qy	421	QNTGWESNANVKGFTQIGAYFCASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480	
Db	438	QVSRQWRKAEVIGTQIGSYFASLCSVDVDTGSTDVLIGAPHYEQTRGGQVSVCP 497	
Qy	481	PRGQARWOCDAVLXGEQGPGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540	
Db	498	PRGWR-RWCDAVLYGEQGHMGRFGAALTVLGDVNGDKLTDVWIGAPGEENRGAVILF 556	
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGHVLLLRSQ 600	
Db	557	HGVLGFSISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVDLVAVARGQVLLLR 616	
Qy	601	PVLRVKAIMEENPREVARVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660	
Db	617	PVLWGVSNQCFPAEIPRAPECREQVVEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676	
Qy	661	YDLALDSGRPHSAFNETKNSTRQTQVLGTQCTETLKLQLPNCIEDPVPSPVILRLNF 720	
Db	677	LDLALAPGRLSPRAIFOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736	
Qy	721	SLVGTPLSAFGLNRPVLAEDAQLFTALPFEKNCNDNICODDLISITFSEMSLDCLVYG 780	
Db	737	TLVGKPELLAFRLNRPMLAALQRYFTASLPFEKNCADHICQDNLGISFSFGLKSLLVG 796	
Qy	781	GPREFNVTVRNDRGDSYRTOVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840	
Db	797	SNLELNAEVMVNDGEDSVGTITTFSHPAGLSRYRYVABGQQLRSLHLTC--CSAPVG 854	
Qy	841	SGALKSTCSINHPIFPENSEVTFTNITFDVDSKASLGNKILLKANVTSENMPRTNKTEF 900	
Db	855	SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVGLDRLLLIANTVSSENNIPTSKTIF 914	
Qy	901	QLELPVKYAVYVTVSHGVSTKVLNETAS-ENTSRVMQHQYQVSNLQSRSLPISLVFLVP 959	
Db	915	QLELPVKYAVYIVVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQGRDLPSVINFVWP 974	
Qy	960	VRNLQVIVDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCQRIQCDD 1019	
Db	975	VELNQAVMNDVEVSHQNPQSLRCSSEKIPAPSDFLAHQKNPVLDCSIAGCLRFECDV 1034	
Qy	1020	PFFGIOEEFNATIKGNLSFDWYIKTSHNLLIVSTAEILPNDVSTFTLLPQOGAFVRSQTE 1079	
Db	1035	PSFSVQBEELDTLKGNLSTFGWVRQILOKQVSVSVSVAEIIIDTSVYSQLPQGEAFMRAQTI 1094	

Qy 1080 TKVEPFEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128

Db 1095 TVLEKYKHNPILPLIVGSSIGLILLALITAVLYKVGFQYKEMMEE 1143

Search completed: November 9, 2004, 12:08:14

Job time : 109.25 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-3
Perfect score: 5879
Sequence: 1 ENLDTENAMTFQENARGFGQ.....FRQYKDMWSEGGPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5879	100.0	1137	10	US-09-902-481A-3
2	5857	99.6	1137	10	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	10	US-09-902-481A-1
5	5839	99.3	1153	10	US-09-891-943-3
6	5839	99.3	1133	14	US-10-144-259-30
7	5839	99.3	1153	14	US-10-207-655-176
8	5836	99.3	1137	10	US-09-902-481A-5
9	5832	99.1	1137	10	US-09-902-481A-6
10	5823.5	99.1	1152	9	US-09-945-265-4
11	3446	58.6	1163	14	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	10	US-09-891-943-4
Sequence 3, Appli					
Sequence 4, Appli					
Sequence 1, Appli					
Sequence 30, Appl					
Sequence 176, App					
Sequence 5, Appli					
Sequence 6, Appli					
Sequence 4, Appli					
Sequence 204, App					
Sequence 4, Appli					

14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	10	US-09-891-943-2
16	3372.5	57.4	1161	9	US-09-350-259-99
17	3372.5	57.4	1161	10	US-09-891-943-99
18	3213.5	54.7	1161	9	US-09-350-259-55
19	3213.5	54.7	1161	10	US-09-891-943-55
20	3201	54.4	1161	9	US-09-350-259-53
21	3201	54.4	1161	10	US-09-891-943-53
22	3192.5	54.3	1151	9	US-09-350-259-37
23	3192.5	54.3	1151	10	US-09-891-943-37
24	3180	54.1	1155	9	US-09-350-259-46
25	3180	54.1	1155	10	US-09-891-943-46
26	1848	31.4	369	13	US-10-087-192-1212
27	1536.5	26.1	1170	9	US-09-945-265-2
28	1532.5	26.1	1170	15	US-10-261-164-1
29	1509	25.7	1223	16	US-10-408-765A-295
30	1341.5	22.8	1086	16	US-10-408-765A-1871
31	1229.5	20.9	494	9	US-09-350-259-103
32	1229.5	20.9	494	10	US-09-891-943-103
33	1128.5	19.2	413	9	US-09-350-259-101
34	1128.5	19.2	413	10	US-09-891-943-101
35	1128	19.2	1179	14	US-10-173-550-2
36	1128	19.2	1179	14	US-10-173-551-2
37	1087	18.5	1188	14	US-10-291-265-810
38	1085.5	18.5	1151	10	US-09-884-130-103
39	1085.5	18.5	1151	10	US-09-836-353A-103
40	1085.5	18.5	1179	10	US-09-918-715-250
41	1085.5	18.5	1179	17	US-10-474-794-250
42	1083	18.4	1188	14	US-10-291-265-338
43	1079.5	18.4	1189	10	US-09-984-130-35
44	1079.5	18.4	1189	10	US-09-836-353A-35
45	1079.5	18.4	1189	15	US-10-262-839-4

ALIGNMENTS

RESULT 1

US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ENLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPT	60
Db	1	ENLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPT	60
Qy	61	RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK	120
Db	61	RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK	120

```
121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVVRELENTNGARKNAFKILFLL 240
181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVVRELENTNGARKNAFKILFLL 240
241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 300
241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 300
301 NFEALKTQONQUREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG 360
301 NFEALKTQONQUREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG 360
361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQVSVCP 480
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQVSVCP 480
481 PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
481 PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
661 YDLALDSGRPHSRAVFNETKNSTRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
661 YDLALDSGRPHSRAVFNETKNSTRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
721 SLVGTPLSAFGLNRPVLVADAOFLTALPFPEKNCNDNICODDLGITFSMSLDCLVVG 780
721 SLVGTPLSAFGLNRPVLVADAOFLTALPFPEKNCNDNICODDLGITFSMSLDCLVVG 780
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVNSLNGORSPLISLVFLPV 960
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVNSLNGORSPLISLVFLPV 960
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PPNVNCISIAVCORIQDIP 1020
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKA PPNVNCISIAVCORIQDIP 1020
1021 FPGIQEFNATLKNLSFDWYIKTSHNHLIVSTAETLFDNDSVFTLLPGOGAFVRSQDET 1080
1021 FPGIQEFNATLKNLSFDWYIKTSHNHLIVSTAETLFDNDSVFTLLPGOGAFVRSQDET 1080
1081 KVEPFEVFNPLPLTVGSSVGLLILALITAALYKLGFFKQYKDMMEGGPPGAEPPQ 1137
1081 KVEPFEVFNPLPLTVGSSVGLLILALITAALYKLGFFKQYKDMMEGGPPGAEPPQ 1137
```

```
Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
; US-09-902-481A-4

Query Match 99.6%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVGAPQEI VAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVGAPQEI VAAANQSGSLYQCDYSTGSCPEI 60
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGICFLFGNLRQPOPK 120
DB 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGICFLFGNLRQPOPK 120
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
QY 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVVRELENTNGARKNAFKILFLL 240
DB 181 RIHFTFKFQNNPNSRLIKPITQLGRTHATGLRKVVRELENTNGARKNAFKILFLL 240
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 300
QY 301 NFEALKTQONQUREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG 360
DB 301 NFEALKTQONQUREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQVSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQVSVCP 480
QY 481 PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRQQRARWQCDVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
DB 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRAVFNETKNSTRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 661 YDLALDSGRPHSRAVFNETKNSTRQTVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
```

```
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 900
Db 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 900
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMOHQYVSNLQORSLSILVELVPV 960
Db 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMOHQYVSNLQORSLSILVELVPV 960
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Qy 1021 FFGIQBEFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLPGQAFVRSQTEP 1080
Db 1021 FFGIQBEFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLPGQAFVRSQTEP 1080
Qy 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGSVVQLQSRVVVGAPQEIIVANQRGSIYQCDYSTGSCPI 60
Db 17 FNLDTENAMTFQENARGFGSVVQLQSRVVVGAPQEIIVANQRGSIYQCDYSTGSCPI 76
Qy 61 RLQVPVEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 180
Db 121 FPALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 180
```

```
Db 137 FPALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTTFEFQNNPNRSLIKPIITQLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTTFEFQNNPNRSLIKPIITQLGRTHRTATGLRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGSKFGDPLGYEDVIPLEDRGVIRVVLGFGDAFRSEKSOELNTVASKPRDHVFOAN 300
Db 257 TDGSKFGDPLGYEDVIPLEDRGVIRVVLGFGDAFRSEKSOELNTVASKPRDHVFOAN 316
Qy 301 NFEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
Db 317 NFEALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRQORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVYLF 540
Db 497 PRQORARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLRLRSQ 616
Qy 601 PVLRVKAIMBPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRLEGIOISVVT 660
Db 617 PVLRVKAIMBPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRLEGIOISVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPSPVTLRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPSPVTLRLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMOHQYVSNLQORSLSILVELVPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMOHQYVSNLQORSLSILVELVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVCNSIAVCQRIQCDIP 1036
Qy 1021 FFGIQBEFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLPGQAFVRSQTEP 1080
Db 1037 FFGIQBEFNATLKNLSFDWIKTSHNHLIVSTAEILFNDVSFTLPGQAFVRSQTEP 1096
Qy 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
```

```
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match          99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRFMKELVSTIMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRFMKELVSTIMEQLKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 240
Db 197 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNIINGARKNAFKILVI 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTIASKPRDHVQVN 316

QY 301 NPEALKTVONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLPSTVGSYDNAG 360
Db 317 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLPSTVGSYDNAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQSVVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQSVVCP 496

QY 481 PRGQARWQCDVLYEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616

QY 601 PVLRVKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
Db 617 PVLRVKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

QY 661 YDLALDSGRPHSAVENETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSAVENETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEADAQRULTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEADAQRULTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPRFNVTVVRNDEGSDYRTQVTFPPFLDLSYRKVSTLQNSORSORSLACESASSTEV 840
Db 797 GPRFNVTVVRNDEGSDYRTQVTFPPFLDLSYRKVSTLQNSORSORSLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFFPENSEVTENITFDVDSKASGLNKLKLLKANVTSENMPRNKTEF 900
Db 857 SGALKSTSCSINHPIFFPENSEVTENITFDVDSKASGLNKLKLLKANVTSENMPRNKTEF 916

QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQGSRLPISLVLVPEV 960
Db 917 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQGSRLPISLVLVPEV 976

QY 961 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPHSDFLAELRKAPVVCNSIAVCORTQCDIP 1020
Db 977 RLNQTVIWDPRQVTFPSENLSSTCHTKERLPHSDFLAELRKAPVVCNSIAVCORTQCDIP 1036

QY 1021 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDVSFTLLPQOGAFVRSQTET 1080
Db 1037 PFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDVSFTLLPQOGAFVRSQTET 1096

QY 1081 KVEPPEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGABPQ 1137
Db 1097 KVEPPEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMMSGGPPGABPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match          99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRFMKELVSTIMEQLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPDPRFMKELVSTIMEQLKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 240
Db 197 RHFTFKFQNNPNRSLIKPTTOLLGRTHATGLRKVVRELFTNIINGARKNAFKILVI 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTIASKPRDHVQVN 316

QY 301 NPEALKTVONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLPSTVGSYDNAG 360
Db 317 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLPSTVGSYDNAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQSVVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQSVVCP 496

QY 481 PRGQARWQCDVLYEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616

QY 601 PVLRVKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
Db 617 PVLRVKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

QY 661 YDLALDSGRPHSAVENETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSAVENETKSNTRTOTVGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
```

```
Db 197 RHFTFKFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRYVLGPDAPFRSEKSRQELMTVASKPRDRHVFOAN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELMTIASKPRDRHVFOVN 316
Qy 301 NFEALKTQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAITNSGPLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAITNSGPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASCLSDVDSNGSTDLVLGAPHYYEOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASCLSDVDSNGSTDLVLGAPHYYEOTRGQGVSCPL 496
Qy 481 PRGORARQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLVRKALMEPNPREVARNVECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 617 PVLVRKALMEPNPREVARNVECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKEF 916
Qy 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNCTVLDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Db 977 RLNCTVLDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1036
Qy 1021 FFGIQEEFNATLKGNSLDFDWIKTSHNHLITVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080
Db 1037 FFGIQEEFNATLKGNSLDFDWIKTSHNHLITVSTAEILFNDVSFTLLPGOGAFVRSOTET 1096
Qy 1081 KVPEFVNPPLPLIVSGSVGGLLLALITAAALYKLGFFKQYKDMGEGGPPGAEPPQ 1137
Db 1097 KVPEFVNPPLPLIVSGSVGGLLLALITAAALYKLGFFKQYKDMGEGGPPGAEPPQ 1153
```

RESULT 6

US-10-144-259-30

; Sequence 30, Application US/10144259

; Publication No. US2003010961A1

; GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin

; APPLICANT: Li, Rui

; APPLICANT: Xiong, Jian-Ping

; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF

```
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.3%; Score 5839; DB 14; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTRYVVKLCFLFGSNLRQOPK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTRYVVKLCFLFGSNLRQOPK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTILFSLMOYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTILFSLMOYSEEF 196
Qy 181 RIHTTFKEFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240
Db 197 RIHTTFKEFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLGYEDVPELDREGVIRYVLGPDAPFRSEKSRQELMTVASKPRDRHVFOAN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELMTIASKPRDRHVFOVN 316
Qy 301 NFEALKTQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAITNSGPLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAITNSGPLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLAGPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASCLSDVDSNGSTDLVLGAPHYYEOTRGQGVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASCLSDVDSNGSTDLVLGAPHYYEOTRGQGVSCPL 496
Qy 481 PRGORARQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLVRKALMEPNPREVARNVECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 617 PVLVRKALMEPNPREVARNVECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTOTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
```

Db 797 GPRENVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQGGAFVRSQTET 1080
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQGGAFVRSQTET 1096
QY 1081 KVEPPEVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1137
Db 1097 KVEPPEVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 14; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYKGLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLDGSIIIPHPRRMKELVSTIMEOLKSKTILFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLDGSIIIPHPRRMKELVSTIMEOLKSKTILFSLMQYSEEF 196
QY 181 RIHFTPKFQNNPNRSLKIPITQLLGRTHATGLARKVRELFTNITNGARKNAFKILFL 240
Db 197 RIHFTPKFQNNPNRSLKIPITQLLGRTHATGLARKVRELFTNITNGARKNAFKILVI 256
QY 241 TDGEKGDPLGYEDVITPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPPEDHVFQAN 300
Db 257 TDGEKGDPLGYEDVITPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPPEDHVFQAN 316
QY 301 NFPAALTKVQNLREKIFAISGTOTGSSSPSEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFPAALTKVQNLREKIFAISGTOTGSSSPSEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILRNVRQSVLIGAPRYQHIGLVAMER 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILRNVRQSVLIGAPRYQHIGLVAMER 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIQTQRCGVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIQTQRCGVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAIVLF 540
Db 497 PRGQARWQCDVAVLYGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAIVLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGSLSGQDLTMDGLVDTLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGSLSGQDLTMDGLVDTLTVGAQGHVLLRSQ 616
QY 601 PVLRYKALMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQKSTRRLRREGQIQSVVT 660
Db 617 PVLRYKALMEFNPREVARNVFECNDQVVKGEAGVRVCLHVQKSTRRLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVENETKNSRRTQTVLGTQTCETLKLQLPNCIEDPVPSPIVRLNF 720
Db 677 YDLALDSGRPHSRAVENETKNSRRTQTVLGTQTCETLKLQLPNCIEDPVPSPIVRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALPPFEKNCNDNICODDLSITFSEMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALPPFEKNCNDNICODDLSITFSEMSLDCLVVG 796
QY 781 GPRENVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 840
Db 797 GPRENVTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 976
QY 961 RLNQTVIWDROPVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQGGAFVRSQTET 1080
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFTLLPQGGAFVRSQTET 1096
QY 1081 KVEPPEVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1137
Db 1097 KVEPPEVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKQKDMSEGPPGABEQ 1153

RESULT 8

US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

US-09-902-481A-5

```
Query Match          99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
Db 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPPRDHVFOAN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPPRDHVFOAN 300

Qy 301 NFEALKTIVONLRREKIPIAEGTGTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTIVONLRREKIPIAEGTGTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480

Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNREGAVYLF 540
Db 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNREGAVYLF 540

Qy 541 HGTSGSGISPSHQSRIAGSKLSPLRLQYFGQSLSGQDLTMDGLVDLTIVGAQGHVLLRSQ 600
Db 541 HGTSGSGISPSHQSRIAGSKLSPLRLQYFGQSLSGQDLTMDGLVDLTIVGAQGHVLLRSQ 600

Qy 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

Qy 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQPNCTEDPVSPIVRLNF 720
Db 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQPNCTEDPVSPIVRLNF 720

Qy 721 SLVGTPLSAGFNLRPVLAEADAQLFTALFFPEKNCGNNDICQDDLSITTFGMSLDCLVVG 780
Db 721 SLVGTPLSAGFNLRPVLAEADAQLFTALFFPEKNCGNNDICQDDLSITTFGMSLDCLVVG 780

Qy 781 GPREFNVTVVRNDGEDSYQTQVTFPPPLLSYHKVSTLQNRQSQRWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDGEDSYQTQVTFPPPLLSYHKVSTLQNRQSQRWRLACESASSTEV 840

Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

Qy 901 QLELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRWQHQQVSNLQORSPLISLVFLVPV 960
Db 901 QLELPVKYAVYVMVTSHGVSSTKYLNFTASNTSRWQHQQVSNLQORSPLISLVFLVPV 960

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDIP 1020
Db 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDIP 1020
```

```
Qy 1021 PFGIOEBFNATLKGNSLSPDWIKTSHNHLILVSTAEIILFNDVSFTLLPQGAFAVSQTET 1080
Db 1021 PFGIOEBFNATLKGNSLSPDWIKTSHNHLILVSTAEIILFNDVSFTLLPQGAFAVSQTET 1080

Qy 1081 KVPEFEVFNPLPLIVGSSVGGLLALITALYKLGFEKROYKDMWSEGGPGABPO 1137
Db 1081 KVPEFEVFNPLPLIVGSSVGGLLALITALYKLGFEKROYKDMWSEGGPGABPO 1137

RESULT 9
US-09-902-481A-6
; Sequence 6 Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6
```

```
Query Match          99.2%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120
Db 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFPGSNLRQOPQK 120

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180
Db 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPPRDHVFOAN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPRSEKSRQELNTVASKPPRDHVFOAN 300

Qy 301 NFEALKTIVONLRREKIPIAEGTGTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTIVONLRREKIPIAEGTGTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL 480

Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNREGAVYLF 540
Db 481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNREGAVYLF 540
```

```
Db 481 PRGQARWQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Qy 541 HGTSGSGISPHSQRIAGSKLSPLRQYFGQSLGGQDLTMGCLVDLTGCAQGHVLLRSQ 600
Db 541 HGTSGSGISPHSQRIAGSKLSPLRQYFGQSLGGQDLTMGCLVDLTGCAQGHVLLRSQ 600
Qy 601 PVLVRKATMEPNPREVARNVFNCDQVVKVKEAGVRVCLHVOKSTRDLRGQIQSVVT 660
Db 601 PVLVRKATMEPNPREVARNVFNCDQVVKVKEAGVRVCLHVOKSTRDLRGQIQSVVT 660
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Db 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Qy 721 SLVGTPLSAFGLRPLVLAEDAQRLETPALPPFKKNGNDNICODDLSITFSFMSLCLVVG 780
Db 721 SLVGTPLSAFGLRPLVLAEDAQRLETPALPPFKKNGNDNICODDLSITFSFMSLCLVVG 780
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLISYRKVSTLQNSRQSWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLISYRKVSTLQNSRQSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
Db 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
Qy 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Qy 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQGAFAVSQTEF 1080
Db 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQGAFAVSQTEF 1080
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137
Db 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shinaka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-945-265-4

Query Match 99.1%; Score 5823.5; DB 9; Length 1152;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCPT 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCPT 76
```

```
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSLNRQOQPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSLNRQOQPK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKELVSTIMEQLKKSKTLFSLMOYSBEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKKSKTLFSLMOYSBEF 196
Qy 181 RIHFTFKFQONPNRSLIKPITQLGRTHATGLRKVVRELFTNTGARKNAFKILFLL 240
Db 197 RIHFTFKFQONPNRSLIKPITQLGRTHATGLRKVVRELFTNTGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGVEDVIPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 300
Db 257 TDGEKFGDPLGVEDVIPELDREGVIRYVLGFDAPRSEKSRQELNTVASKPRDHVFOAN 316
Qy 301 NFEALKTQVONLREKIFAIEGTQTGSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQVONLREKIFAIEGTQTGSSSPHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTBVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMER 420
Db 377 GVFLYTSKEKSTFINMTBVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMER 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRG-QARWQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
Qy 541 HGTSGSGISPHSQRIAGSKLSPLRQYFGQSLGGQDLTMGCLVDLTGCAQGHVLLRSQ 600
Db 556 HGTSGSGISPHSQRIAGSKLSPLRQYFGQSLGGQDLTMGCLVDLTGCAQGHVLLRSQ 615
Qy 601 PVLVRKATMEPNPREVARNVFNCDQVVKVKEAGVRVCLHVOKSTRDLRGQIQSVVT 660
Db 616 PVLVRKATMEPNPREVARNVFNCDQVVKVKEAGVRVCLHVOKSTRDLRGQIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 735
Qy 721 SLVGTPLSAFGLRPLVLAEDAQRLETPALPPFKKNGNDNICODDLSITFSFMSLCLVVG 780
Db 736 SLVGTPLSAFGLRPLVLAEDAQRLETPALPPFKKNGNDNICODDLSITFSFMSLCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLISYRKVSTLQNSRQSWRLACESASSTEV 840
Db 796 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLISYRKVSTLQNSRQSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 975
Qy 961 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 976 RLNQTVINDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1035
Qy 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQGAFAVSQTEF 1080
Db 1036 PFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQGAFAVSQTEF 1095
Qy 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137
Db 1096 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1152
```

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 58.6%; Score 3446; DB 14; Length 1163;
Best Local Similarity 60.5%; Pred. No. 1.2e-284;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY 1 FNLDTENAMTQENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCBPI 60
Db 20 FNLDTENAMTQENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCBPI 79
QY 61 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFPGSNLRQOPQK 120
Db 80 GLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFPGSNLRQOPQK 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEF 180
Db 138 LPVSRQECPRQEQDIFVLIDGSGSISSRNFMFVRAVISQRPSTQSLMQFSNKF 197
QY 181 RIHTTFEFQNNPRSLIKIPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILFL 240
Db 198 QTHPTFEFRSTNPLSILASVHQLQGYTATATQNVVHRLPHASYGARDAKILIVI 257
QY 241 TDGKFGDPLGYEDVPELDRGVIRYVGLGDAFRSEKSEQLNTYASKPRDHVQAN 300
Db 258 TDGKFGDPLGYEDVPELDRGVIRYVGLGDAFRSEKSEQLNTYASKPRDHVQAN 317
QY 301 NFEALKTQVQNLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSTVGSYDAG 360
Db 318 DFDALKDQNLQEKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSTVGSYDAG 377
QY 361 GVELYTSKEKSTFIMTRVDSMDNDAVLGYAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 378 GAFILYPPNMGPTFIMNSQENVRDSDYLGYSTELALWKGVSQSLVGLAPRYOHTKAVIFT 437
QY 421 QNTGMESNANVKTQIGVAGSLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSCPL 480
Db 438 QVSRQWMAEVTGTQIGVAGSLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSCPL 497
QY 481 PRQARWQCDVILYGRQGFQWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGQGHQWGRFGAALTVLGDVNGDKLTDVVGAPGEENRGAYVLF 556
QY 541 HGTSGSGISPSHSORVAGSLSPRLQYFGSLGGQDLTMDGLVLTGAGQHVLLRSQ 600
Db 557 HGVLPISPSHSORVAGSLSPRLQYFGSLGGQDLTMDGLVLTGAGQHVLLRSQ 616
QY 601 PVLVRKAIMFNPREVARNFECDNVKGEAGEVRVCHVQKSTRDLRBEQIQSVVT 660
Db 617 PVLVGVGVMQFIPAEIPRSAFEQEVVSEQTLVQSNICLYIDKRSKNLLGSRDQSSVT 676

QY 661 YDLALDSGRPHSRVAFNETKSTRBQTQVGLGTQTCETLKLQIPNCIEDPVSFIVLRNF 720
Db 677 LDLALDPGRLSPRATFOETKNSRLSRVRLGLKAHCENFNLLPSCVEDSVTPITRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPFPEKNCNNDNICODDLSITSEFSLDCLVVG 780
Db 737 TLVGRPLLAFLNRPVLAEDAQRFLTALFPFPEKNCNNDNICODDLSITSEFSLDCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYKRVSTLQNRQSRWELACESASSTEV 840
Db 797 SNLELNAEVMVMWNGEDSYGTTITFSPAGLSYRYVAEGQKQGLRSLHLTCDSPVGV-- 854
QY 841 SGALKSTSCSINHPIFFENSEVTFNITFDVDSKASLGNKLLKANVTSENMMRINKTEF 900
Db 855 SQGTWSTSCRINHPIFFENSEVTFNITFDVDSKASLGNKLLKANVTSENMMRINKTEF 914
QY 901 QLELPVYAVVYVTVSHGVSTKYLNFTAS-ENTSRVMOHOYQVSNLQORSLSPISELVLP 959
Db 915 QLELPVYAVVYVTVSHGVSTKYLNFTAS-ENTSRVMOHOYQVSNLQORSLSPISELVLP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHDSFLAELKAPVNVCSIAVCQIQCDI 1019
Db 975 VELNQEAVMMDVEVSHPONPSLRCSSEKIAAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034
QY 1020 PFFGIQEEFNATLKNLSFDFWIKTSHNHLIIVSTAELLFNDVSFTLLPGQAFVRSOTE 1079
Db 1035 PSFVQBELDFTLKNLSFDFWIKTSHNHLIIVSTAELLFNDVSFTLLPGQAFVRSOTE 1094
QY 1080 TKVPEFVNPPLIIVGSSVGGILLALITAAALYKLGFFKRYQKMMSE 1128
Db 1095 TVLEKYVHNPTPLIVSSIGILLALITAAALYKLGFFKRYQKMMSE 1143

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.2%; Score 3423; DB 9; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTQENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCBPI 60
Db 20 FNLDTENAMTQENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCBPI 79
QY 61 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFPGSNLRQOPQK 120
Db 80 GLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQCSENTYVKGCLFPGSNLRQOPQK 137

```
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSSEF 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSSEF 197
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVVRELFNTNGARKNAFKILFL 240
Db 198 QTHFTFEEPRRTSNPLSLASVHQOGFTYTATQNVVHRLFHASYGARRDAIKILIVI 257
QY 241 TDGKFGDPLGVEDVPELDRGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFOAN 300
Db 258 TDGKFGDPLGVEDVPELDRGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFOAN 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTSSSFEHMQSGFSAITNGPLSTVGSYDWAG 360
Db 318 DFDALKDIONQKKEKIFAIEGTQTSSSFEHMQSGFSAITNGPLSTVGSYDWAG 377
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGLVAMFR 437
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTOIGSFYFASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLYGEOQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEOQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGQHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGQHVLLRSQ 616
QY 601 PVLVRKAIMFNPREVARNVFCNDQVWKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVSMQFIPAEIPRSAFECEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSFIVLELNF 720
```

RESULT 13
US-09-891-943-4

; Sequence 4, Application US/09891943

```
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.2%; Score 3423; DB 10; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGGQSVQVQCGSRVVVGAPQEIIVANQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTENAMTFOENARGGQSVQVQCGSRVVVGAPQEIIVANQORSLYQCDYSTGSCPEI 79
QY 61 RLOVPEAVNMISLGLSLAATTPSPOLLACGPTVHQCSTNTYVKGLCFFGSLNLRQPOK 120
Db 80 GLQVPEAVNMISLGLSLAATTPSPOLLACGPTVHQCSTNTYVKGLCFFGSLNLRQPOK 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSSEF 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSSEF 197
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVVRELFNTNGARKNAFKILFL 240
Db 198 QTHFTFEEPRRTSNPLSLASVHQOGFTYTATQNVVHRLFHASYGARRDAIKILIVI 257
QY 241 TDGKFGDPLGVEDVPELDRGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFOAN 300
Db 258 TDGKFGDPLGVEDVPELDRGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFOAN 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTSSSFEHMQSGFSAITNGPLSTVGSYDWAG 360
Db 318 DFDALKDIONQKKEKIFAIEGTQTSSSFEHMQSGFSAITNGPLSTVGSYDWAG 377
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGLVAMFR 437
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTOIGSFYFASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLYGEOQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEOQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGQHVLLRSQ 600
Db 557 HGVLGPSISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGQHVLLRSQ 616
QY 601 PVLVRKAIMFNPREVARNVFCNDQVWKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVSMQFIPAEIPRSAFECEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSFIVLELNF 720
```

Db 677 LDALAPGRSLSPRAIFQETKNSLSRVVLGLKAHCENFNLLPSCVEDSVIPIILRLNF 736
Qy 721 SILVGTPLSAGNLRVLAEDAQRLFTALFPPEKNCNNDNI CODDLSITFSFMSLDCILVVG 780
Db 737 TLVGKPLLAFLRNLRLPMLAALAQRYFTASLPPEKNGADHI CODNLIGISFSPGLKSLVG 796
Qy 781 GPRFNVTVTVRNDCGEDSYRTQVTFEFLDLISYKVVSTLQNRORSWRLACESASSTEV 840
Db 797 SNLELNAEVMWVNDGEDSYGTTITESHAGLSYRVVARGQKQQLRSLLHLC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVPSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVVMVTSYHGVSTKYLNFTAS-ENTSRVMOHQVQVSNLQORSPIPSLVFLVP 959
Db 915 QLELPVKYAVVIVSSHQFTKYLNFSSEKESHVAMHYQVNNLQORDLFSVINEFWP 974
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDI 1019
Db 975 VELNQEAVMWVESHQPNLSRCSSEKIAPPASDFLAHQNPVLDSCSIAGCLRFRCDV 1034
Qy 1020 PFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
Db 1035 PSFSVQEBLDTLKNLSFGWVRQILQKVSWSVAEIIIDTSVVSQLPQGEAFMRAQTI 1094
Qy 1080 TKVPEFVNPPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMXSE 1128
Db 1095 TVLEKIKVHNPIPLIVGSSIGLGLLILALITAVLYKVGFFKQYKEMMBE 1143

RESULT 14

US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 57.6%; Score 3388; DB 9; Length 1161;
Best Local Similarity 58.9%; Pred. No. 1,1e-279;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTFQENARFGGSVQVLOGSRVVGAPQFIVAANQRGSLYQCDYGTSGCEPI 60
Db 17 FNLDVEEPTIFQEDAGFGGSVVGFGSRLVVGAPLEVAANQTRGLYDCAAAATGMCQPI 76
Qy 61 RLQVPVEAVNMSGLSLAATSPQILACGPTHQTCSENTYUKGLCFLFGSLNRQOPQK 120
Db 77 PLHIRPAVNMSGLTLAATNSGRLLACGPTHVRVCGENSYSKSGSCLLLGSRW-ELIQT 135

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKELVSTIMEOLKSKTILFSLMOYSBEF 180
Db 136 VPDATPECPHQEMDIVFLIDGSGSDQDFNQMGFVQAVMGQFEGTDTTLFALMOYSNLL 195
Qy 181 RIHFTTFKEFONNPNRSLIKPITOLLGRTHATGLRKVVRELNFINTNGARKNAPKILFLL 240
Db 196 KIHFTTFOFRSPQOQSLVDPIVLKGLTFTATGILTVVTOLFHHKNGARKSAKKILLIV 255
Qy 241 TDGSKFGDPLGYEDVIPELDREGVIRVVLGFGDAPRSEKSRQELNVTASKPPDRHVFQAN 300
Db 256 TDGQYKIDPLEYSDVIPAQKAGIIRYAIGVGHAFQGTARQELNVTISSAPPQDHVFQVD 315
Qy 301 NFEALKTVQNLREKIFALECTQTGSSSPHEMSOEGFSAAITSNGLLSTVGSYDWAG 360
Db 316 NFAALGSIQOLQKIKIYAVEGTQSRASSSPHEMSOEGFSAITMDGGLFLGAVGSFWSG 375
Qy 361 GVFLYTSKEKSTFINMTVRDSDMMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLAVMFR 420
Db 376 GAFLYPPNMSPTFINMQENVDMEDSVLYGYSTELALMKGVQNLVLGAPRYOHTGKAVFT 435
Qy 421 QNTGWSNSNANVKGTQIGAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQVSVCP 480
Db 436 QVSRQWRKKAEBVTGTQIGSYFGASLCSDVDSDGSTDLILIGAPHYYEQTRGGQVSVCP 495
Qy 481 PRGORAEWQCDVLYGQGGPWGFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVLYF 540
Db 496 PRGORVQWQCDVLYGQGGPWGFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVLYF 555
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYFGOSLGGQDLTMDGLVDTLTVGAOGHVLILRSQ 600
Db 556 HGASESGISPSHSORIASQSLSPRLQYFGQALSGQDLTQDGLMDLAVARGQVLLRLSL 615
Qy 601 PVLVKAINFNPREVARNVPCNDQVVKGEAGEVRVCLHVQKSTRDRRREBQIQSVVT 660
Db 616 PVLKGVAMRFPSPVEVAKAVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
Qy 661 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQENCLDEDPVSVILNLF 720
Db 674 FDLALDPGRLTSTRAIFNETKNPTLTRRKTILGLHCHETLKLDPDCEVDVSPILHLNF 733
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFLTALFPPEKNCNNDNI CODDLSITFSFMSLDCILVVG 780
Db 734 SLVREPIPSQNLRPVLAQSGQDLFTASLPPEKNCQGDGICEGDLGVTLSFGSLQTLTVG 793
Qy 781 GPRENVTVTVRNDCGEDSYRTQVTFEFLDLISYKVVSTLQNRORSWRLACESASSTEV 840
Db 794 SSLELNVITVMNAGEDSYGTVWSLYYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVPSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 853 EG-LRSCRSVNHPIFHEGNGTFTVTFDVSYKATLGDRLMLRASSENKASSSKATF 911
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVSNLQORSPIPSLVFLVP 959
Db 912 QLELPVKYAVVITMSRQEESTKYFNATSDKKMKAEHRYRVNNLSQORDLAISINFWP 971
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDI 1019
Db 972 VLLNGVAVMDVWMEAPSQSL--PCVSEKPPQHSDFLTQISRPFMLDCSIADCLQPRCDV 1029
Qy 1020 PFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
Db 1030 PSFSVQEBLDTLKNLSFGWVRQILQKVSWSVAEIIIDTSVVSQLPQGEAFMRAQME 1089
Qy 1080 TKVPEFVNPPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMXSE 1128
Db 1090 MVLEEDVEVNAIPIINGSSVGLLILALITATLYKLGFFKRYHKEMLEED 1138

```

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-2

Query Match      57.6%; Score 3388; DB 10; Length 1161;
Best Local Similarity 58.9%; Pred. No. 1,1e-279;
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVLQGGRRVVVGAPOEIVAAQNRQSLYOCDYSGSCPEI 60
DB 17 FNLDEEPTIFQEDAGGFGQSVVQFGGRLVVGALEVVAAQNRQSLYOCDYSGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATSPPOILLACGPTVHOTCSNTVYKGLCFGLSGNLQQPOK 120
DB 77 PLHIREAVNMSLGLTAASTNGSRLACGPTLHRRVCGNSYSKSGCLLGSRW-EIQT 135
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELIVSTIMEQLKKKTLFSLMOYSEEF 180
DB 136 VPDATPECPHOEMDIVFLIDGSGSIDQDNFNQMGFVQAVMGQFEGTDTLFAIMOYSNLL 195
QY 181 RHFTFKFQNNPNRSLIKPTOLLGRTHATGLKVVRELFNITNGARKNAFLIFLL 240
DB 196 KIHFTTQPTSPSQSLVDPIVLQKGLTFTATGILTVTQLFHHKNGARKAKKILIVI 255
QY 241 TDGKFGDPLGVEDVITPELDRSGVTRYVLGFGDAPRESEKSRQBLNTVASKPRDHFVOAN 300
DB 256 TDGQKYKDPLEYSDVLPQEKAGIIRYALGVGHAFQGPARGELNITSSAPQDHFVKVD 315
QY 301 NFEALKTQNLREKIFALEGTQTSSESSFEHMSQEGFSAATSNGLPSTVGSYDWAG 360
DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 375
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAILRNVRQSLVLGAPRYOHIGLVAMER 420
DB 376 GAFLYPPNMSPTFINNSQENVMDRDSYLGSTELALWKGVQNLVLGAPRYOHTGKAVIF 435
QY 421 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 480
DB 436 QVSRQWRKAEVTGTQIGYFGASLCSVDVDSGSTDLLIGAPHYVEOTRGQVSVCP 495
QY 481 PRGQARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 496 PRGQVQWQCDVLRGEQHPMGRFGAALTVLGDVNDKLDVAIGAPGEQNRGAVYLF 555
QY 541 HGTSGSGISPSHSQRIAGSLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVILLRSQ 600
DB 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLRSL 615
QY 601 PVLVRKAIMFENPREAVNFCNDQVKGKEAGEVRVCLHVQKSTRDLRREGIOQSVVT 660
DB 616 PVLKVGWAMFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
QY 661 YDLALDGRPHSRVAVNETKNSRRTQVGLTQTCTETKLQLPNCIEDPVSPIVLRNLF 720

```

```

DB 674 FDALDPCRLTSRAIFNETKNPTLTRRKTGLGHCETKLKLLPDCVEDVVSPILHLNF 733
QY 721 SLVGTPLSAFCNLRPVLAEADAQRLFTALFPPEKNCNDNIQDOLSIETSEFMSLDCLVVG 780
DB 734 SLVREPSPQNLRPVLAVGSDLTASLFPPEKNCQDCLCEGLGVTLSFSGLTIVG 793
QY 781 GPRBFNTVTVVRNDGEDSYRTQVTFPFFPLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840
DB 794 SSLELNVITVYVWNAAGEDSYGVVSLIYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRINKTEF 900
DB 853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTPDVSYKATLGDRLMRASASSENKASSKATF 911
QY 901 QLELPVKYAVYVWVTSHGVSSTKYNF-TASENTSRVMOHQVSNLQORSIPISLVFLVP 959
DB 912 QLELPVKYAVYVWVTSHGVSSTKYNF-TASENTSRVMOHQVSNLQORSIPISLVFLVP 971
QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVWVNSIAVCQRIQDI 1019
DB 972 VLLNGVAVWDVYVMEAPSQSL-PCVSEKPKPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1029
QY 1020 PFFGIQEEFNATLKNLSFSDWIKTSHNHLILVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
DB 1030 PSFSVQEEBLDTLKNLSFGWVRETQKKVLVWSVAEITFDTSVYSQPLPGQCAFVRAQME 1089
QY 1080 TKVEPPEVFNPLPLIVGSSVGLLALLITLTAALYKLGFFKRYKQKMMSE 1128
DB 1090 MVLEDEEVYNAIPIITGSSVGNALLLITLTAALYKLGFFKRYKQKMMSE 1138

```

Search completed: November 9, 2004, 12:46:48
Job time : 102.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMWSGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A-COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B-COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A-COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B-COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/FACTUS-COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5839	99.3	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5839	99.3	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5839	99.3	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5839	99.3	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5839	99.3	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5839	99.3	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5839	99.3	1153	3 US-08-193-043-3	Sequence 3, Appli
9	5839	99.3	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5839	99.3	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5808.5	98.8	1152	2 US-08-476-062A-43	Sequence 43, Appl
12	5808.5	98.8	1152	5 PCT-US96-01314-43	Sequence 43, Appl
13	5808.5	98.8	1152	6 5424399-2	Patent No. 5424399
14	3446	58.6	1163	2 US-08-476-062A-44	Sequence 44, Appl
15	3446	58.6	1163	5 PCT-US96-01314-44	Sequence 44, Appl
16	3423	58.2	1163	1 US-08-173-497-4	Sequence 4, Appli
17	3423	58.2	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3423	58.2	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3423	58.2	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3423	58.2	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3423	58.2	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3423	58.2	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3423	58.2	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3423	58.2	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3423	58.2	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3388	57.6	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3388	57.6	1161	1 US-08-286-889-2	Sequence 2, Appli

28 3388 57.6 1161 1 US-08-485-618-2 Sequence 2, Appli
29 3388 57.6 1161 1 US-08-362-652-2 Sequence 2, Appli
30 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli
31 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli
32 3388 57.6 1161 2 US-08-943-363-2 Sequence 2, Appli
33 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli
34 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli
35 3388 57.6 1161 4 US-09-350-259-2 Sequence 2, Appli
36 3372.5 57.4 1161 1 US-08-485-618-99 Sequence 99, Appl
37 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appl
38 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appl
39 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appl
40 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appl
41 3372.5 57.4 1161 4 US-09-688-307A-99 Sequence 99, Appl
42 3372.5 57.4 1161 4 US-09-350-259-99 Sequence 99, Appl
43 3213.5 54.7 1161 3 US-09-193-043-55 Sequence 55, Appl
44 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appl
45 3213.5 54.7 1161 4 US-09-350-259-55 Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
|||||
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEARLGCQPDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEARLGCQPDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 316
QY 301 NFEALKTQNLQREKIFAEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKSTFINNTRVDSMDNAYLGYAAAIIILNRVQSLVLCAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKSTFINNTRVDSMDNAYLGYAAAIIILNRVQSLVLCAPRYOHIGLVAMFR 436
QY 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
DB 437 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 496
QY 481 PRQARWOCDAVLXGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQARWOCDAVLXGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HTSGSGISPSHSQRIAGSKLSPRIQYFQOSLSGGQDLTMDGLVLDLTGCAQHVLLRSQ 600
DB 557 HTSGSGISPSHSQRIAGSKLSPRIQYFQOSLSGGQDLTMDGLVLDLTGCAQHVLLRSQ 616
QY 601 PVLVRKAIEMEPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIEMEPREVARNVFECNDQVKGKAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAFVNETKNSRTRQTVLGLTQTCETILKQLPNCIEDPVPVILRLNF 720
DB 677 YDLALDSGRPHSAFVNETKNSRTRQTVLGLTQTCETILKQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGNLRPVLAEDAQLFTALPPFEKNCNDNICODDLSITFSMSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEDAQLFTALPPFEKNCNDNICODDLSITFSMSLDCLVVG 796
QY 781 GPREFNVTVRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNSORSQSWRLACESASSTEV 840
DB 797 GPREFNVTVRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNSORSQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 916
QY 901 QLELPKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSIPTSLVFLPV 960
DB 917 QLELPKYAVYVMTSHGVSTKYLNFTASENTSRVHQYQVSNLQORSIPTSLVFLPV 976
QY 961 RLNQTVIDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORICODIP 1020
DB 977 RLNQTVIDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORICODIP 1036
QY 1021 FFGIQEENATLKNLSFDWYIKTSHNHLIVSTABEILFNDVSFTLLPGOGAFVRQSTET 1080
DB 1037 FFGIQEENATLKNLSFDWYIKTSHNHLIVSTABEILFNDVSFTLLPGOGAFVRQSTET 1096
QY 1081 KVEPFEPVNPPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMMSGGPPCAEPQ 1137
DB 1097 KVEPFEPVNPPLIVGSSVGLLILLALITAAALYKLGFFKQYKDMMSGGPPCAEPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/082866889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEARLGCQPDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEARLGCQPDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILFL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 316
QY 301 NFEALKTQNLQREKIFAEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAEGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 376


```
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGPYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHRAVFNKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720
Db 677 YDLALDSGRPHRAVFNKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPPEKNGNDNICODDLSTIFSMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPPEKNGNDNICODDLSTIFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNRQSRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 900
Db 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTET 916
Qy 901 QLELPVKYAVVMVTSHGVSPTKYNFTASENTSRVMOHQVSNLQSRSLPISLVELVPV 960
Db 917 QLELPVKYAVVMVTSHGVSPTKYNFTASENTSRVMOHQVSNLQSRSLPISLVELVPV 976
Qy 961 RLNTQVWRDPQVTFSENLSSTCHTKERLPKSHSDFLAELKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNTQVWRDPQVTFSENLSSTCHTKERLPKSHSDFLAELKAPVNVCSIAVCQRIQCDIP 1036
Qy 1021 FFGQIESFNATLKNLSFDWIKTSHNELLIVSTAEILFNDVSVTLPLPGQAFVRSQTET 1080
Db 1037 FFGQIESFNATLKNLSFDWIKTSHNELLIVSTAEILFNDVSVTLPLPGQAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLPLTVGSSVGLLLALITAAALYKLGFFKRYKDMSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPLTVGSSVGLLLALITAAALYKLGFFKRYKDMSEGGPPGAEPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3
```

```
Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLGSRVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHTCSENTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHTCSENTYVKGCLFLFGSNLRQPOK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTLPSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTLPSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHTATGLRKVVRELFNITNGARKNAFKILL 256
Qy 241 TDGEKFGDPLGYEDVIPELDREGVIRVYVVGDAFRSEKRGKQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVIPELDREGVIRVYVVGDAFRSEKRGKQELNTVASKPPRDHVFQAN 316
Qy 301 NFPAKTVQNLREKIPAEIGTQTGSSSSSEHEMSQGFSAAITNSGPELLSTVGSYDWAG 360
Db 317 NFPAKTVQNLREKIPAEIGTQTGSSSSSEHEMSQGFSAAITNSGPELLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGPYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
```

557 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGQDLTMDGLVLTGVAQHVLRLRSQ 616
601 PVLVRKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLVRKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQTCETLKLQLPNCIEDPSPVILRLNF 720
677 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQTCETLKLQLPNCIEDPSPVILRLNF 736
721 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQSWRLACESASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQSWRLACESASSTEV 856
841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
901 QLELPPKAVYVMVTSHGVSSTKYNFTASENTSRVWQHGYQVSNLQSRSLPISLVFLVFPV 960
917 QLELPPKAVYVMVTSHGVSSTKYNFTASENTSRVWQHGYQVSNLQSRSLPISLVFLVFPV 976
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPDSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPDSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
1021 PFGIOBEFNATLKNLSFPMYIKTSNHLIYVSTAEILFNDSEVFTLLPCGAFVRSQDET 1080
1037 PFGIOBEFNATLKNLSFPMYIKTSNHLIYVSTAEILFNDSEVFTLLPCGAFVRSQDET 1096
1081 KYEPFEPVNPPLVLGVSSVGGILLALITAALYKLGFFRQYKDMWSEGGPPGAEPPQ 1137
1097 KYEPFEPVNPPLVLGVSSVGGILLALITAALYKLGFFRQYKDMWSEGGPPGAEPPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQVQSGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQVQSGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFPGSNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFPGSNLRQPOK 136

QY 121 PPEALRGCPQEDSDTAFLIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLIMOYSSEF 180
DB 137 PPEALRGCPQEDSDTAFLIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLIMOYSSEF 196

QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAKILFLL 240
DB 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAKILFLL 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVTRYVVGFDGAFRSKSRQELMTVASKPRDHFVQAN 300
DB 257 TDGEKFGDPLGYEDVIPLEDRGVTRYVVGFDGAFRSKSRQELMTVASKPRDHFVQAN 316

QY 301 NFEALKTQVQNLREKIFAETGTQTSSESFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTQVQNLREKIFAETGTQTSSESFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLI GAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLI GAPHYYEOTRGQVSVCP 496

QY 481 PRGQARWQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTVAIGAPGEENRGAVYLF 540
DB 497 PRGQARWQCDVLYGEGQPGWRFGAALTVLGDVNGDKLTVAIGAPGEENRGAVYLF 556

QY 541 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGQDLTMDGLVLTGVAQHVLRLRSQ 600
DB 557 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGQDLTMDGLVLTGVAQHVLRLRSQ 616

QY 601 PVLVRKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLVRKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQTCETLKLQLPNCIEDPSPVILRLNF 736

QY 721 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQSWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPPLDLISYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLISLFLVPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLISLFLVPV 976
Qy 961 RLNQTVIWDNRQVTFSENLSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCIP 1020
Db 977 RLNQTVIWDNRQVTFSENLSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCIP 1036
Qy 1021 PFGIQEERNATLKNLSDWIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEF 1080
Db 1037 PFGIQEERNATLKNLSDWIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEF 1096
Qy 1081 KVEPPEVENPLPLIVGSSVGGILLIALITALYKLGFFKROVKDMWSEGGPPGAEPQ 1137
Db 1097 KVEPPEVENPLPLIVGSSVGGILLIALITALYKLGFFKROVKDMWSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605.672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ENLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAORGLSYQCDYSTSCSCEPI 60
Db 17 ENLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAORGLSYQCDYSTSCSCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLSGNLQQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLSGNLQQPOK 136
Qy 121 PPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKELVSTIMEQLKKSKTFLSMOYSEEF 180
Db 137 PPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKELVSTIMEQLKKSKTFLSMOYSEEF 196
Qy 181 RIHFTFEKFNPNPRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILVVI 240
Db 197 RIHFTFEKFNPNPRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVIPELDREGVIRYVVGFDGAFRSEKSEQLNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVIPELDREGVIRYVVGFDGAFRSEKSEQLNTVASKPPRDHVFQAN 316
Qy 301 NFEALKTVONOLREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGPELLSTVSGSYDWAG 360
Db 317 NFEALKTVONOLREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGPELLSTVSGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGYAAAIIILNRVQSLVGLAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKCTQIGAFYGASLCSVDVDSNGSTDLVLIGAPHYEQTEGGOVSVCP 480
Db 437 QNTGMWESNANVKCTQIGAFYGASLCSVDVDSNGSTDLVLIGAPHYEQTEGGOVSVCP 496
Qy 481 PRGQRAWQCDVLYGEGQGWGRFGAALTIVLGVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGQRAWQCDVLYGEGQGWGRFGAALTIVLGVNGDKLTDVAIGAPGEDNREGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTWGAQHVLILRSQ 600
Db 557 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTWGAQHVLILRSQ 616
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETHKSTRQTQVGLTQTCETLKLQPNCTEDPVPSPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNETHKSTRQTQVGLTQTCETLKLQPNCTEDPVPSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFEKNCNDNICODDLSITFMSLDCVLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFEKNCNDNICODDLSITFMSLDCVLVVG 796
Qy 781 GPREFNVTVVRNDGEDSVRTQVTFPPPLDLISYRKVSTLQNRQSRWELACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPPLDLISYRKVSTLQNRQSRWELACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLISLFLVPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSPLISLFLVPV 976
Qy 961 RLNQTVIWDNRQVTFSENLSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCIP 1020
Db 977 RLNQTVIWDNRQVTFSENLSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCIP 1036

QY 1021 PFGIOEENATIKGNLSFDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFRSQET 1080
DB 1037 PFGIOEENATIKGNLSFDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFRSQET 1096
QY 1081 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGABPQ 1137
DB 1097 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGABPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPEIIVAAORGSLYQCDYSTGSCPT 60
DB 17 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPEIIVAAORGSLYQCDYSTGSCPT 76
QY 61 RLQVPEAVNMISGLSLAATSPQLIACGPTVHQCSENTRYVYKGLCFLFGSNLRQOPQK 120
DB 77 RLQVPEAVNMISGLSLAATSPQLIACGPTVHQCSENTRYVYKGLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTFLSLMQSYSEF 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTFLSLMQSYSEF 196
QY 191 RIHFTFKFQONPNPRSLIKITQLLGRTHATGLRKVVRELFINITGNARKNAFKILFLL 240
DB 197 RIHFTFKFQONPNPRSLIKITQLLGRTHATGLRKVVRELFINITGNARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVGLFGDAPRSEKSRQELMTVASKPRDHVFOAN 300
DB 257 TDGEKFGDPLGVEDVPELDRGVIRYVGLFGDAPRSEKSRQELMTVASKPRDHVFOAN 316
QY 301 NFEALKTVQNOLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSVTGSDYDAG 360
DB 317 NFEALKTVQNOLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSVTGSDYDAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMER 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVVCP 496
QY 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540
DB 497 PRGQARWQCDVAVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLOYFGOSLGGQDLTMDGLVDLTIVCAQGHVLLLRQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLOYFGOSLGGQDLTMDGLVDLTIVCAQGHVLLLRQ 616
QY 601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660
DB 617 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQPCIEDPVPVILRLNP 720
DB 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQPCIEDPVPVILRLNP 736
QY 721 SLVGTPLSAFNLPRVLAEDAOQLFTALPPFEKNGCNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNLPRVLAEDAOQLFTALPPFEKNGCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLGYRKVSTLQONORSQSWLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLGYRKVSTLQONORSQSWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLAFTASENTRVMOHOYQVSNLQORSILSLVLPV 960
DB 917 QLELPVKYAVYVMTSHGVSTKYLAFTASENTRVMOHOYQVSNLQORSILSLVLPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPHSDFLAELKAPVNVNCSTAVCORIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPHSDFLAELKAPVNVNCSTAVCORIQCDIP 1036
QY 1021 PFGIOEENATIKGNLSFDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFRSQET 1080
DB 1037 PFGIOEENATIKGNLSFDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFRSQET 1096
QY 1081 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGABPQ 1137
DB 1097 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROYKDMMSGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.3%, Score 5839; DB 2; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQEIIVANQKGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQEIIVANQKGLYQCDYSTGSCPEI 76
QY 61 RLQVPVAVNMSGLSLAATTPPQLAACGPTVHQTCSNTYVKGCLCFLGSLNLRQOPQK 120
Db 77 RLQVPVAVNMSGLSLAATTPPQLAACGPTVHQTCSNTYVKGCLCFLGSLNLRQOPQK 136
QY 121 FPEALRGCPDESDIAFLDGSGSIIPDPRMKELVSTIMEQLKSKTLFSLMQXSEEF 180
Db 137 FPEALRGCPDESDIAFLDGSGSIIPDPRMKELVSTIMEQLKSKTLFSLMQXSEEF 196
QY 181 RIHFTKFEFQNNPRSLIKPIQLLGRTHATGLRKVRELNTINGARKNAFKILVLI 240
Db 197 RIHFTKFEFQNNPRSLIKPIQLLGRTHATGLRKVRELNTINGARKNAFKILVLI 256
QY 241 TDGEKFDPLGYEDVPEADREGVIRVVLGFGDAFRSEKSKQLNTVASKPPRDHVFQAN 300
Db 257 TDGEKFDPLGYEDVPEADREGVIRVVLGFGDAFRSEKSKQLNTVASKPPRDHVFQAN 316
QY 301 NFALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATSNGLPILSTVGSYDWAG 360
Db 317 NFALKTVQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATSNGLPILSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMER 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLGAPHYEQTGTGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLGAPHYEQTGTGGQSVCP 496
QY 481 PRGQARWQCDVILYGEQCPWGRFGAALTVLGVDVNGDKLTDAIAGPEENRGAIVLF 540
Db 497 PRGQARWQCDVILYGEQCPWGRFGAALTVLGVDVNGDKLTDAIAGPEENRGAIVLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKATMEFNPREVARNVEFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKATMEFNPREVARNVEFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRRTQVGLGTOTCETLKLQLPNCIEDPSPVIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRRTQVGLGTOTCETLKLQLPNCIEDPSPVIVLRNF 736
QY 721 SLVGTPLSAFONLRPVLAEADAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFONLRPVLAEADAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKRVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHQYQVSNLQORSIPISLVFLVPV 976
QY 961 RLNQTVIWDPRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSVFTLLPGOGAFVRSOTET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSVFTLLPGOGAFVRSOTET 1096
QY 1081 KVEPFVEVPNPLIIVGSSVGLLALLALITAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
Db 1097 KVEPFVEVPNPLIIVGSSVGLLALLALITAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

```
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match      99.3%; Score 5839; DB 3; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLFSLMOYSEEF 196

QY 181 RHFTTFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RHFTTFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVILGFGDAFRSEKSRQELNVTASKPPRDHVFQ 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVILGFGDAFRSEKSRQELNVTASKPPRDHVFQ 316

QY 301 NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLTSKESKFINNTRVDSMDNAYLGVAAILNRVQSLVILGAPYQHIGLVAMPR 420
DB 377 GVFLTSKESKFINNTRVDSMDNAYLGVAAILNRVQSLVILGAPYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAFYASGLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAFYASGLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 496

QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVILF 540
DB 497 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVILF 556

QY 541 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGQAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGQAQGHVLLRSQ 616

QY 601 PVLVRKAIMFENPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
DB 617 PVLVRKAIMFENPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSIVLRINF 720
DB 677 YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSIVLRINF 736

QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALPFEKNCGNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFNLRPVLAEDAQLFTALPFEKNCGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLVSRYKYSTLQNRQSRQSLACSSASSTEV 840
DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLVSRYKYSTLQNRQSRQSLACSSASSTEV 856

QY 841 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGKLLKANVTSENMBRTKNTF 900
DB 857 SGALKSTSCSINHPIIPENSEVFNITFDVDSKASLGKLLKANVTSENMBRTKNTF 916

QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHQVSNLQORSILPISILVFLPV 960
```

```
RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 278667/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-307A-3
```

```
Query Match      99.3%; Score 5839; DB 4; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFSGNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLFSLMOYSEEF 196

QY 181 RHFTTFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
DB 197 RHFTTFKEFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNITNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVILGFGDAFRSEKSRQELNVTASKPPRDHVFQ 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVILGFGDAFRSEKSRQELNVTASKPPRDHVFQ 316

QY 301 NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
```

Db 317 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKGITGAYFGASCLSDVDNSGSDTLVLGAPHYIEQTRGGQVSVCP 480
Db 437 QNTGWSNANVKGITGAYFGASCLSDVDNSGSDTLVLGAPHYIEQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRYKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRYKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTQVGLGTQTCETLKLQPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTQVGLGTQTCETLKLQPNCIEDPVPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 916
Qy 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHOYVSNLQORSLSILVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHOYVSNLQORSLSILVFLVPV 976
Qy 961 RLNQTIVWRDPQVTFSENLSSTCHTKERLPKSHSDPFLAELRKAPVVCNSIAVCQRIQCDIP 1020
Db 977 RLNQTIVWRDPQVTFSENLSSTCHTKERLPKSHSDPFLAELRKAPVVCNSIAVCQRIQCDIP 1036
Qy 1021 PFGIQEFENATLKNLSPDWIKVSHNHLILVSTABILFNDVSFTLLPGQGAFFVRSQDET 1080
Db 1037 PFGIQEFENATLKNLSPDWIKVSHNHLILVSTABILFNDVSFTLLPGQGAFFVRSQDET 1096
Qy 1081 KVEPFEPVNPPLIVGSSVGLLALLITALYKLGFFKRYKQKDMSEGGPGBAEPQ 1137
Db 1097 KVEPFEPVNPPLIVGSSVGLLALLITALYKLGFFKRYKQKDMSEGGPGBAEPQ 1153

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 4; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQORSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQORSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTQCSNTYVKGCLFGLGSLNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTQCSNTYVKGCLFGLGSLNLRQPOK 136
Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPDPRMKELVSTIMEOLKSKTILSLMOYSEEP 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPDPRMKELVSTIMEOLKSKTILSLMOYSEEP 196
Qy 181 RIHFTFEFQNNPNSRLIKEITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTFEFQNNPNSRLIKEITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 256
Qy 241 TDGKFGDPLGYEDVIVELDREGVIRVYVGLGDAFPRSEKSOELNTVASKPPDRHVFQAN 300
Db 257 TDGKFGDPLGYEDVIVELDREGVIRVYVGLGDAFPRSEKSOELNTVASKPPDRHVFQAN 316
Qy 301 NFEALKTVQNLREKIFAIEGTQGTSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 360
Db 317 NFEALKTVQNLREKIFAIEGTQGTSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSILVLCAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKGITGAYFGASCLSDVDNSGSDTLVLGAPHYIEQTRGGQVSVCP 480
Db 437 QNTGWSNANVKGITGAYFGASCLSDVDNSGSDTLVLGAPHYIEQTRGGQVSVCP 496
Qy 481 PRGORARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRYKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRYKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAVNETKSTRQTQVGLGTQTCETLKLQPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTQVGLGTQTCETLKLQPNCIEDPVPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQORSORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSOTET 1080
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSOTET 1096
QY 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137
Db 1097 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1153

RESULT 11

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275

GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPQEIIVAAANQRGSLVQCYSTYSGSCBPI 60
Db 17 FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPQEIIVAAANQRGSLVQCYSTYSGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGTVHTQTCSENTYVXGLCLFLGSLNLRQQPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGTVHTQTCSENTYVXGLCLFLGSLNLRQQPQK 136
QY 121 FPEALRGCPQSDIDIAELIDGSGSIIIPDFFRMKELVSTIMBOLKSKSTLFLSLMYSEEF 180
Db 137 FPEALRGCPQSDIDIAELIDGSGSIIIPDFFRMKEFVSTVMEQLKSKSTLFLSLMYSEEF 196
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGRLKRVVRELFNTNGARKNAFKILVVI 240
Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGRLKRVVRELFNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGDADFSEKSKROELNTVASKPRDHVFOAN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGDADFSEKSKROELNTVASKPRDHVFOAN 316
QY 301 NPEALKTVQNLREKIFAIEGTQTGSSSSSFHEMSQEGFSAAITSNGPLLTSTVGSYDNAG 360
Db 317 NPEALKTVQNLREKIFAIEGTQTGSSSSSFHEMSQEGFSAAITSNGPLLTSTVGSYDNAG 376
QY 361 GYFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHI GLVAMPR 420
Db 377 GYFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHI GLVAMPR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYETRGQVSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYETRGQVSVCP 496
QY 481 PRGQARWOCDAVLYGEOGPWGRFGAALTVDLVGVDGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWOCDAVLYGEOGPWGRFGAALTVDLVGVDGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSISPSHSQRIAGSKLSPLRQYFQOSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSQRIAGSKLSPLRQYFQOSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 615
QY 601 PVLRVKAIMERNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSPRDLRREGQIOSVVT 660
Db 616 PVLRVKAIMERNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSPRDLRREGQIOSVVT 675
QY 661 YDLALDSDGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPIVLRNLF 720
Db 676 YDLALDSDGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPIVLRNLF 735
QY 721 SLVGTPLSAFNLRPVLAEDAORLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNLRPVLAEDAORLFTALPPFEKNCGNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTYVRNDGEDSYRTQVTFPPLDLSYRKVSTIQNORSORSWRLACESASTEV 840
Db 796 GPREFNVTYVRNDGEDSYRTQVTFPPLDLSYRKVSTIQNORSORSWRLACESASTEV 855
QY 841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 915
QY 901 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 975
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSTCHTKERLPHSDFLAELRKAPVNVCSIAVCORIQCDIP 1035
QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSOTET 1080
Db 1036 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSOTET 1095
QY 1081 KVEPEVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137


```

Db 1096 KVEPEVENPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1152
|||||
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 98.8%; Score 5808.5; DB 5; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMVGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLCFGLGNSLRQOPQK 120
Db 77 RLQVPVAVNMVGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLCFGLGNSLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTTMEQLKKGKTTFLSMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKELVSTTMEQLKKGKTTFLSMQYSEEF 196
Qy 181 RIHFTFEFQNNPNRSLKPIITQLGRTHATGLRVVRELNITNGARKNAFKILFL 240
Db 197 RIHFTFEFQNNPNRSLKPIITQLGRTHATGLRVVRELNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPIGYEDVPELDEGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPIGYEDVPELDEGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQVN 316
Qy 301 NFALKTVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAATTSNGPLISTVGSYDWAG 360
Db 317 NFALKTVQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAATTSNGPLISTVGSYDWAG 376

```

```

Qy 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTROGGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTROGGQSVCP 496
Qy 481 PRGRARWQCDVILYGEQGPWGRFGAALTIVLGDVNGDKLTDVAITGAPGEEDNRGAVILF 540
Db 497 PRG-RARWQCDVILYGEQGPWGRFGAALTIVLGDVNGDKLTDVAITGAPGEEDNRGAVILF 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAAQCHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAAQCHVLLRSQ 615
Qy 601 PVLRVKALMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 616 PVLRVKALMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 675
Qy 661 YDLALDSGRPHSRAVFNETKNSTRRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSRAVFNETKNSTRRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
Qy 721 SLVGTPLSAFNLPRVLAEDAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNLPRVLAEDAQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTIVRNDGEDSYRTQVTFPPDLDSYKVTSLONORSORSWRLACESASSTEV 840
Db 796 GPRESNVTIVRNDGEDSYRTQVTFPPDLDSYKVTSLONORSORSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVFNTITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVFNTITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYINFTASENTSRVMOHQYQVSNLQORSUPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYINFTASENTSRVMOHQYQVSNLQORSUPISLVFLVPV 975
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
Db 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1035
Qy 1021 FFGIQEFNATLKNLSFDWYIKITSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIQEFNATLKNLSFDWYIKITSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPFEVNPPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1152

```

```

RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2
; LENGTH: 1152
5424399-2

```

Query Match 98.8%; Score 5808.5; DB 6; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;									
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;									
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	76						
QY	61	RLOVPVEAVNMSLGLSLAATTSPPQLLAGCTVHTQCSNTYVVKGLCLFLGSLNRLOQPOK	120						
Db	77	RLOVPVEAVNMSLGLSLAATTSPPQLLAGCTVHTQCSNTYVVKGLCLFLGSLNRLOQPOK	136						
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTILFSLMOYSEEF	180						
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTILFSLMOYSEEF	196						
QY	181	RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL	240						
Db	197	RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL	256						
QY	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGFGDAFRSEKSRQELNTVASKPRDHVFQAN	300						
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGFGDAFRSEKSRQELNTVASKPRDHVFQAN	316						
QY	301	NFEALKTQONLREKIFALEGTQGTSSSFEHMQEGFSAITNSGPIILSTVGSYDWAG	360						
Db	317	NFEALKTQONLREKIFALEGTQGTSSSFEHMQEGFSAITNSGPIILSTVGSYDWAG	376						
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGLVAMFR	420						
Db	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGLVAMFR	436						
QY	421	QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGSTDVILGAPHYETRGQSVVCP	480						
Db	437	QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGSTDVILGAPHYETRGQSVVCP	496						
QY	481	PRGQARWQCDVLYGEQOQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540						
Db	497	PRGQARWQCDVLYGEQOQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	555						
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGQAQHVLLRSQ	600						
Db	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGQAQHVLLRSQ	615						
QY	601	PVLRUKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGIOQSVVT	660						
Db	616	PVLRUKAIMEFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRREGIOQSVVT	675						
QY	661	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF	720						
Db	676	YDLALDSGRPHSRVAFNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF	735						
QY	721	SLVGTPLSAFNLPRVLAEDAQRLFTALPPFEKNCNDNIQODDLGITFSFMSLDCLVVG	780						
Db	736	SLVGTPLSAFNLPRVLAEDAQRLFTALPPFEKNCNDNIQODDLGITFSFMSLDCLVVG	795						
QY	781	GPREFNVTVVRNDGDSYRTQVTFEFPDLSYRKVSTLQONQSRQSWRLACESASSTEV	840						
Db	796	GPREFNVTVVRNDGDSYRTQVTFEFPDLSYRKVSTLQONQSRQSWRLACESASSTEV	855						
QY	841	SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900						
Db	856	SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	915						
QY	901	QLELPVKYAVYVMVTSHGVTCKYLNFTASENTSRVMOHQYVNSLNGORSPIISLVFLPV	960						
Db	916	QLELPVKYAVYVMVTSHGVTCKYLNFTASENTSRVMOHQYVNSLNGORSPIISLVFLPV	975						
QY	961	RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQDIP	1020						
Db	976	RLNQTIVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQDIP	1035						
QY	1021	FFGIQEBFNATLKGNSLDFDWMYIKTSHNHLIVSTABILFNDVSTLLPGQAFVRSQTE	1080						

Query Match 58.6%; Score 3446; DB 2; Length 1163;
Best Local Similarity 60.5%; Pred. No. 5.1e-278;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI	60						
Db	20	FNLDTBELTAFRVDSAGFGDSVVQYANSWVVGAPQKIITAANTGGLYQCGYSTGACEPI	79						
QY	61	RLOVPVEAVNMSLGLSLAATTSPPQLLAGCTVHTQCSNTYVVKGLCLFLGSLNRLOQPOK	120						
Db	80	GLQVPPEAVNMSLGLSLASTTSQLLACGPTVHHEGGRNMYLTGLCFLGPT--OULTQR	137						
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTILFSLMOYSEEF	180						
Db	138	LPVSRQECPEQEDIVFLIDGSGSISSRNFEATMNFVRAVISQFQRPSTQFSLMQFSNKF	197						
QY	181	RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL	240						

Search completed: November 9, 2004, 12:20:59
Job time : 28 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481B-4
Perfect score: 5884
Sequence: 1 FNLPTENAMTFQENARGFGQ.....FKROYKDMVSGGPPGAEPPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 RWHULB	cell surface glyco
2	4460	75.8	1153	2 S00551	leukocyte surface
3	3469	59.0	1163	1 RWHULC	cell surface glyco
4	1543.5	26.2	1170	2 S03308	cell surface glyco
5	1527.5	26.0	1163	2 I56126	lymphocyte fuction
6	1140	19.4	1179	2 A53213	integrin alpha-E c
7	1093.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45314	integrin alpha-1 c
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1055	17.9	1180	2 A35854	integrin alpha-1 c
11	1054	17.9	1181	2 A33998	integrin alpha-2 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.4	1041	2 T31437	alphanap integrin -
16	579.5	9.8	1034	2 JC7294	cell surface glyco
17	572.5	9.7	1051	2 A35761	VLA-3 alpha subuni
18	567.5	9.6	1053	2 I55534	integrin alpha-5 c
19	555.5	9.4	1053	2 S4250	integrin alpha-v c
20	546.5	9.3	1034	2 A36108	integrin alpha-6 c
21	535	9.1	1044	2 T10050	integrin alpha-6 c
22	534.5	9.1	1072	2 A38457	integrin alpha-v c
23	532	9.0	1049	2 A27079	integrin alpha-6 c
24	532	9.0	1073	2 B36429	fibronectin recept
25	530.5	9.0	1051	2 A40021	integrin alpha-6 c
26	526	8.9	1048	2 A47421	integrin alpha-5 c
27	525.5	8.9	1091	2 A21543	integrin alpha-6 c
28	517	8.8	1044	2 I6516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

ALIGNMENTS

RESULT 1

RWHULB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M; eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C

B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307146

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965,'P',967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594

A:Note: the authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was confir

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes

A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A:Reference number: A32218; MUID:89098993; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HTC>

A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

integrin alpha-2b
position-specific
integrin alpha-7 c
alpha-7 integrin -
glycoprotein IIB -
leukocyte adhesion
F54F2.1 protein -
integrin alpha v c
hypothetical prote
integrin alpha cha
glycoprotein IIB -
integrin alpha cha
glycoprotein IIB -
integrin alpha-1 -

30 495.5 8.4 1039 2 A34269
31 494 8.4 1394 2 A29637
32 489 8.3 1137 2 JC5950
33 488 8.3 1135 2 I61186
34 486.5 8.3 1037 2 A60163
35 486 8.3 126 2 B30892
36 470 8.0 1106 2 S38783
37 462 7.9 1226 2 S44824
38 453 7.7 1045 2 S60571
39 444.5 7.6 1139 2 S28277
40 426 7.2 1115 2 T09403
41 421.5 7.2 1115 2 T09433
42 391 6.6 764 2 I36916
43 309.5 5.3 1086 2 T18523
44 299 5.1 604 2 I36917
45 286.5 4.9 272 2 A55348

A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-1
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FLE>
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaut, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: 152567; MUID:92144986; PMID:1346576
A;Accession: 152567
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219
A;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: Promoter contains a GATA motif and two Sp1 consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;593-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQFNARGFGQSVVQLGSRVVGAPQEIIVAAQNGSLGYOCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQFNARGFGQSVVQLGSRVVGAPQEIIVAAQNGSLGYOCDYSTGSCPEI 76
QY 61 RLQQPVVEAVNMSLGLSIAATTSPQLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 120
Db 77 RLQQPVVEAVNMSLGLSIAATTSPQLACGPTVHQTCSENTYVKGCLFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLDGSGSIIPHFRNKEWVSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLDGSGSIIPHFRNKEWVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RHFTKFEQNNPNRSLIKPIPTQLGRTHATGLKRVRELFNITNGARKNAFKILFL 240
Db 197 RHFTKFEQNNPNRSLIKPIPTQLGRTHATGLKRVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVDAFSEKSRQELNTVASKPPRDHFVQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFSEKSRQELNTVASKPPRDHFVQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFAAITSNGLPLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFAAITSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYOHIGLVAMER 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLGAPRYOHIGLVAMER 436
QY 421 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVSCPL 496
QY 481 PRGQARWQCDVAVLYGEGQGPWGRFGAALTIVLGVNWDGKLTVDVAIGAPGEDNRGAVILP 540
Db 497 PRGQARWQCDVAVLYGEGQGPWGRFGAALTIVLGVNWDGKLTVDVAIGAPGEDNRGAVILP 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPFEKKGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPFEKKGNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPRFENVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRORSORSLACESASSTEV 840
Db 797 GPRFENVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRORSORSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVSITKYLNFASNTSRVMQHQYQVSNLQSRSLPISLFLVLPV 960
Db 917 QLELPVKYAVVMVTSHGVSITKYLNFASNTSRVMQHQYQVSNLQSRSLPISLFLVLPV 976
QY 961 RLNQTVIWDROPVTFSENLSSCTHFKERLPSHSDFLAELRKAPVWNCSTAVQCRIQCDDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSCTHFKERLPSHSDFLAELRKAPVWNCSTAVQCRIQCDDIP 1036
QY 1021 FFGIOEEFNATLKGNSLDFWYIKTSHNHLIIYSTAEILFNDSVFTLLPQGGAFVRSQTT 1080
Db 1037 FFGIOEEFNATLKGNSLDFWYIKTSHNHLIIYSTAEILFNDSVFTLLPQGGAFVRSQTT 1096
QY 1081 KVEPEFENVPLPIVGVSSVGGLLLALITAAIYKLGFFKROVKDMWSEGGPPCAEPQ 1137
Db 1097 KVEPEFENVPLPIVGVSSVGGLLLALITAAIYKLGFFKROVKDMWSEGGPPCAEPQ 1153
RESULT 2
S00551
leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse
N;Alternate names: complement-3 receptor alpha chain
C;Species: Mus musculus (house mouse)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C;Accession: S00551; 159078
R;Pytel, R.
EMBO J. 7, 1371-1378, 1988
A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the
A;Reference number: S00551; MUID:88312584; PMID:3044779
A;Accession: S00551
A;Molecule type: DNA
A;Residues: 1-1153 <PYT>
A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G529
A;Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: 159078; MUID:86287312; PMID:2942940
A:Accession: 159078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <WAA>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;
Best Local Similarity 73.8%; Pred. No. 2.4e-302;
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFOENARGQGVVQLGSRVVGAPQEIIVANQRGSLVQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFOENAKGFGQNVVLGTSVVVAAPOEAKAVNQTGALVQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTYKGLCLFLGSLNRQPOK 120
Db 77 PLQVPPEAVNMSLGLSLAVTSVPOQLLACGPTVHONCKENTYVNGLCYLFGLSNLRPPQQ 136

QY 121 FPAALRGCPQEDSDIAPLIDGSGSIIPHPFRMKEWSTVMEQLKSKTFLSLMOXSEEF 180
Db 137 FPEALRECPQESDIAPLIDGSGSINNIDFQMKFVSTVMEQFKSKTFLSLMOXSEEF 196

QY 181 RIHFTKEFQNNPNSLIPIQLLGRTHATGLKRVBELFNITNGARKNAFKILFL 240
Db 197 RIHFTFNDFRNPSHSHVSPKQLNGRTKTASGIRKVRVRELFHKTNGARENAKILVVI 256

QY 241 TDGEKFGDPLGYEDVTPELDREGVIRVIGVGAERSEKRSQRLNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYKDVTPEDRAGVIRVIGVGNFKNPQSRRELDTIASKPGEHFVQVD 316

QY 301 NFEALTIQNLREKIPIAEGTQTSGSSFEHMSQEGFSAATISNGPLLSITVSGVSDWAG 360
Db 317 NFEALTIQNLREKIPIAEGTQTSGSSFEHMSQEGFSAATISNGPLLSITVSGVSDWAG 376

QY 361 GVFLYTSKESKTFINTRVDSMDNDAYLGAAIILRNVRQSLVLGAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTRVDSMDNDAYLGASAVILNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQSVCP 480
Db 437 ENFGTWEPTSIGSQIGSYFGASLCSVDMDADGNTNLILIGAPHYVEKTRGQSVCP 496

QY 481 PRGQARWQCDVLYGEOQGPWRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 540
Db 497 PRG-RARWQCEALLHGDQHPWRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 555

QY 541 HGTSGSGISPSHSQRTAGSKLSPRIYFQOSLSGGQDLTMDGLMDLTVGAQGHVLLRSQ 600
Db 556 YGASIASLSASHRIIGAHFSGLYFQOSLSGGKDLTMDGLMDLTVGAQGHVLLRSQ 615

QY 601 PVLVRKAIEMFNPREVARNVFCNDQVKGKAGEVRVCLHVOKSTFDRLRGQTSVVT 660
Db 616 PVLRLAETWTFEPKVKVARSVFACQEOVLKNKQAGEVRVCLVRKNTKDLREGDTQSTVT 675

QY 661 YDLALDSGRPHSRAVNETKNSRRQTQVLGHTQTCETLKLQPLNCIEPVPPIVLRNF 720
Db 676 YDLALDPVRSIRAFDETKNTRRTQVGLMQRCETLKLILPCVDVDSVPILRLNY 735

QY 721 SLVGTPLSAGNLRVLAEDAQRLLFTALPPFEKNGCNDDI CODDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLRPLVLAEDAQRFFTFAMPPFEKNGCNDSICODDLSITWSAMGLDLVVG 795

QY 781 GPRBNVTVTVRNDGDSRYRTQVTFPPFLDLISYRKVSTLQNRQSRL-ACESASSTE 839
Db 796 GPQDFNMSVTLRNDGDSYGTQVTVYVPSGLSYKDSASQNLTKKFWFKPAPRESSSE 855

QY 840 VSGALKSTCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTE 899
Db 856 GHGALKSTTNWNIHPIIPANSEVTFNITFDVDSHAFGNKLLKAIIVASENNMGRTHTK 915

QY 900 FQLELPVKYAVMVVTSGVSTKYNFTASENTSRVMQHOYQVSNLQSRSLPISIVFLVP 959
Db 916 FQLELPVKYALYMIYVTSDESSIRVNLNTASEMTSKVHQHQFNNGQSLVSVVFWIP 975

QY 960 VRLNQTWIDRPQVTFENISSTCHTKERLPSHSDFLAELRKAPVAVNCIAVACORIQCDI 1019
Db 976 VQINNVTVMDHPQVIFSONLSSACHTEQKSPHNSFRDQLERTFVLNCSVAVCKRIQCDL 1035

QY 1020 PFFGIQEEFNATLGNLSFDWYIKTNSHLLIYVSTAILNDSVFTLLPGGAFVRSQTE 1079
Db 1036 PSFNTQEIFNVTLKGNLSFDWYIKTNSHLLIYVSTAILNDSAFALLPGQESVRSKTE 1095

QY 1080 TKVEPEFVPNPLPIVGVSSVGGLLILALITAAALYKLGFEKRYKDMSEGPPGAEPQ 1137
Db 1096 TKVEPEVHNVPVPLIVGVSSIGGLVLLALITRAGLYKLGFRQYKDMNEAAPQAPQ 1153

RESULT 3

RWHLIC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A:Residues: 1-1163 <COR>

A:Cross-references: UNIPROT:P20702

A>Note: this revision to the sequence from reference A35543 includes the carboxyl end

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.

A:Reference number: A35543; MUID:90153906; PMID:2303426

A:Accession: A35543

A:Molecule type: DNA

A:Residues: 1-834 <CO2>

A>Note: this sequence has been revised in reference A36584

R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A:Reference number: S00864; MUID:88166645; PMID:3327687

A:Accession: S00864

A:Molecule type: mRNA

A:Residues: 1-755,'L',757-1163 <CO3>

A:Cross-references: GB:M81695; EMBL:Y00093; NID:9487829; PIDN:AAA59180.1; PID:g487830

A>Note: part of this sequence was confirmed by protein sequencing

C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my

C:Genetics:

A:Gene: GDB:ITGAX; CD11C

A:Cross-references: GDB:119758; OMIM:151510

A:Map position: 16p11.2-16p11.2

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo

C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <SIG>

F:20-1107/Domain: extracellular #status predicted <EXT>

F:149-319/Domain: von Willebrand factor type A repeat homology <WAA>

F:1108-1133/Domain: transmembrane #status predicted <TM>

F:1134-1163/Domain: intracellular #status predicted <INT>

F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match	59.0%	Score 3469;	DB 1;	Length 1163;	
Best Local Similarity	60.9%	Pred. No. 3.1e-233;			
Matches	688;	Conservative 142;	Mismatches 293;	Indels 6;	Gaps 4;
QY	1	FNLDTENAMTFQ-ENARGFGQSVVQLQSGSRVVVVGAPQEIIVAAQNGSLYQCDYGTGSCPEI	60		
Db	20	FNLDTEELTAFRVDSAGFGSVQVYANSWVVGAPQKITAAQNTGGIYQCGYGTGACEPI	79		
QY	61	RLQVPVEAVNMSLGLSLAAATTPSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQK	120		
Db	80	GLQVPVEAVNMSLGLSLASITSPQLLACGPTVHHEGGRNMYLTGLCFLLGPT--QLTOR	137		
QY	121	FPFALRCPOEDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQKSKTFLPSLMQYSEEP	180		
Db	138	LPVSRQCPQEQDIIIVELIDGSGSISRNPFATMNFVRAVISQFQRPSTQFSLMQFKNKF	197		
QY	181	RIHFTPFQFNPNRSLIKPIITCLLGRTHPTATGLRKVVRELFINITGARKNAFKILFL	240		
Db	198	QTHFTFEFRRTSNPLSILLASVHQLQGFYTTATAIQNVVHRLFASVGARDDATKILIVI	257		
QY	241	TDGKFGDPLGYEDVIBELDREGVIRYVIGVGDAPRSEKSEQLNTVASKPRDHVFOIN	300		
Db	258	TDGKKGSLDYKDVPMDADAAGIIRYAGVGLAFQNRNSWKELNDIASPSQEHIPKVE	317		
QY	301	NFEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAITNSGPELLSTGVSYDWAG	360		
Db	318	DFDALKDIONLREKIFAIEGTQIGSFGASLCSVDVTDGSDTLVLGAPHYEQTRGGQSVVCP	377		
QY	361	GVFLYTSKEKSTFNMTREVDSDMDAYGLGYAAAILNRVQSLVILGAPRYQHIGLVAMFR	420		
Db	378	GAFIYPPNMSPTFINMSQENVMDRSDYLGISTELALWKGVQSLVILGAPRYOHTKAVIFT	437		
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLGAPHYEQTRGGQSVVCP	480		
Db	438	QVSRQWRMAEVTGTQIGSFGASLCSVDVTDGSDTLVLGAPHYEQTRGGQSVVCP	497		
QY	481	PRGORARQCDVLYGEGQGPWRFGAALTIVLGVNGDKLTDVAIGARDEENRGAVYLF	540		
Db	498	PRGNR--RWVCDVLYGEGQGPWRFGAALTIVLGVNGDKLTDVWIGAFGEENRGAVYLF	556		
QY	541	HGTSSGSLSPHSQRIAGSKLSPRIQYFGQSLGGQDLTMGCLVDLTVGAQHVLLLRSQ	600		
Db	557	HGVLGSLSPHSQRIAGSKLSRLQYFGQSLGGQDLTMGCLVDLTVGAQHVLLLRQ	616		
QY	601	PVLRVYKAIEMFNPREVARNVFCNDQVYVKGKAGEVRVCLVHVKQSTRDLRRLRGQTQSVVT	660		
Db	617	PVLWTVGSMQFIPAEIPRSAPFECQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676		
QY	661	YDLALDSGRPHRAVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRNF	720		
Db	677	LDLALDPGLSPRATFOETKNSLSRVRVLGKAHCENFNLLPLPSCEVDSVPTITRLNF	736		
QY	721	SLAVGTPLSAFNLRPVLAEDAORLTALPPEKNCNDNICODDLISITFSFMSLCLVVG	780		
Db	737	TLVHGKPLAFLRNLRLMLAADAGRYFTASLPFNKCGADHICQDNLGISFPGLSLAVG	796		
QY	781	GPREFNVTVVRNDEGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWSWELACESASTVB	840		
Db	797	SNLELNAEVMWVNDGSDSYGTTITFSPHAGLSYRYVAEQKQQLRSLHLTCDSPAVG--	854		
QY	841	SGALKSTCSINHPIFENPSEVFNITFDVDSKASIGNKLLKANVTSENMPRTNKTFF	900		
Db	855	SGQWTSTSCRINHLIPFGAQITFLATFDVSPKAVIGDRLLLTANVSSNNPTRTSKTTF	914		
QY	901	QLELPVKYAVMYVTSHGVSTKYLNFTAS--ENTSRVMOHQYQVNSLQGRSLFISLVFLVP	959		
Db	915	QLELPVKYAVTVVSSHEQFTKYLNFSSEKESHVAMHYQVNNLQGRDLPSVINFVWP	974		
QY	960	VRINQTVWDRPOVTSSENLSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORQCDI	1019		
Db	975	VELNQEAVMVDVESHQPNPNSLSCSEKIAAPPASDFLAHIQKNPVLDCSIAGCLRPRCDV	1034		

1020 PFFGIQEFNATLKNLSFDWYIKTSHNHLIIVSTAEILLPNDVSFTLLPQCGAFVRSOTE 1079

1035 PFSVQOEELDFTLKGNLSPGWVRQIILKKVSVVVAEITPDTSVSQLPQGEAFMRAQTT 1094

1080 TKVPEPEVENPPLIIVGSSVGGLLALITAAALYKLGFFKRYKDMWSE 1128

1095 TVLEKYKHVHTPLIIVGSSITGGLLALITAVLYKVGFFKRYKEMWEE 1143

RESULT 4

S03308

cell surface glycoprotein CD11a precursor - human

N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 09-Jul-2004

C;Accession: S03308; A47458; A47565; A48759; S36044

J;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.

J;Cell Biol. 108, 703-712, 1989

A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit

A;Reference number: S03308; MUID:89139587; PMID:2537322

A;Accession: S03308

A;Molecule type: mRNA

A;Residues: 1-1170 <L>

A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UEC8; EMBL:X00796; NID:G31421; PIDN:CAA68

A;Note: part of this sequence was confirmed by protein sequencing

R;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.

Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993

A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a)

A;Reference number: A47458; MUID:93248261; PMID:8097887

A;Accession: A47458

A;Molecule type: DNA

A;Residues: 1-20 <C>

A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)

R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993

A;Title: Identification of cell-specific and developmentally regulated nuclear factors

A;Reference number: A47565; MUID:93281759; PMID:8099450

A;Accession: A47565

A;Molecule type: DNA

A;Residues: 1-20 <SHE>

A;Cross-references: GB:M95609

R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.

J. Biol. Chem. 268, 19305-19311, 1993

A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.

A;Reference number: A48759; MUID:93374910; PMID:8103515

A;Accession: A48759

A;Molecule type: DNA

A;Residues: 1-20 <NUE>

A;Cross-references: EMBL:Z22804; NID:G311405; PIDN:CAA80461.1; PID:G311406

C;Genetics:

A;Gene: GDB:ITGAL; CD11A

A;Cross-references: GDB:119757; OMIM:153370

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom

C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr

F;1-25/Domain: signal sequence #status predicted <SIG>

F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted

F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.2%; Score 1543.5; DB 2; Length 1170;

Best Local Similarity 34.3%; Pred. No. 6.2e-99; Indels 101; Gaps 37;

Matches 403; Conservative 210; Mismatches 460;

QY 1 FNLDTENAMTFQ--ENARGFGQSVVQLQSGSRVVVVGAPQEIIVAAQNGSLYQCDYGTGSCSE 58

Db 26 YNLVDVGARSFPFPRAGRHFGRVQLV--GNGVIVGAPGE---GNSTGSLYQCSGTGHCL 81

QY 59 PRLQVPVEAVNMSLGLSLAAATTPSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLR--- 115

Db 82 PVTLR--GSNYSKYLGMTLATDPTDGSILACDPGLSRITDQNTYLSGLCYLFPQNLQGP 140

QY 116 -QOPQKFPALACPOEDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQKSKTFLPSLM 174

141 LQGRPGFQCEIKG----NVDVLVFLPDGMSLQDPDEFQKILDFKMDVMKKLSNTSYQFAAV 196
175 QYSEEFRIIHTFKEFQNNPNRSLIKPITQLGRTHATATGLKRVVRELENIINGARKNAF 234
197 QFSTSYKTFDFSDYVVKMDPDALLKHVKHMLLTNTFGAINVATEVFEELGARPDPAT 256
235 KILFLLTDCGKGDPLGYEDVTPELDREGVIRVVGDAFSEKSRQBELNTVASKPPRD 294
257 KVLIIITDGE--ATDSGNTDAAD-----IIRYIIGIKHFQTKESQETLHKFASKPASE 309
295 HVFQINNFEALKTIONLREKIFALETGTGSSSPHEHMSQEGSAAITNSGPLLSTVG 354
310 FVKILDTPEKLDLTELQKLYIVIEGTSKQDLSFNMLSSSGISADLSRGHAVVAVG 369
355 SYDMAGGVF-LVTSKEKSTFTINMTRVDSMDNDAYLGYAAA-IILNRVQSLVIGAPRQ 412
370 AKDMAGGFLDLKADLQDQDTFIGNELTPEVRAGLYGTVTWLPSQKTSLLASGAPRQ 429
413 IGLVAMER--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVIGAPHYRQT 470
430 MERVLLFQEPQGGHMSQVTHGTQIGSYFGGELCGVDVDQDGETELLIGAPLFYGEQ 489
471 RGGQSVSCPLPRGQARWQDAV--LYGEOGQPMGRFGAALTIVLGDVNGDKLTDVAIGAP 528
490 RGRVFIY----QRRQLGFVEVSELDQDPGLRFGAALTALTDINGDGLVDVAVGAP 544
529 GEEDNRGAVLYPHGTSGSGISPSHSQRJAGSKLSPLQYFGQSGGQDLTMDGLVDLTV 588
545 LEE--QGAVYIFNGRHG-GLSPQSPQRIEQTQVLSGIQMFGRSHGVKDLBGDGLADVAV 601
589 GAQGHVILLRSQPLRVKAINMEFNPREVARNVFECDQV--KGKAGEVRVCLHVQKSTR 647
602 GAESQMVLSRPVVDVMVTLMSFSPAETIPVHEVECSYSTSNKMEGVNITICFOI-KSLY 660
648 DRLREGIQSVVTVYDLALDSGRPHSRVAVENETKSTRQTVLGTQTCETLKLQLPNCI 707
661 PQF-QGRVLNAVLTYLQDGHTRRRGLFPGRHHLRRIAVT-TSMSCDTDFSFHFVVCV 718
708 EDVPSPIVLRNFSL----VGTPLS--AFGN-----LRPVLAEADAQRLFTALFPEKNCGN 757
719 QDILSPINVSINLSLWEEGTFRDQRAQGDIPILRPSLHSETWEI-----PPEKNCGE 773
758 DNICODLSLTFSPMSLDCLVGGPREPNVTVVRNDEGDSYRQTVTFPPDLISYRKVS 817
774 DKCEANLRVFSFARSRALELTAFASLSVELSLNLEEDAYWYQLDHFPPGLSPKVE 833
818 TLQNRQSQRWRLACES--ASSTEVSGALKSTCSINHPIFPENSEVTFNITFDVSKAS 875
834 ML---KPHSQIPVSCHELPRSLRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887
876 LGNKLILKANVTSENN----MPRTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFNTASEN 931
888 WGDVSELHANVTENNEDSDLESNATII---IPILYPINILQDQEDSTLYVSFTPKGP 944
932 TSRVMOHOYQV---SNLQORSLP-LSLVFLVPLVRLNQTVIWDPRQVTFSENLSSTCHTK- 986
945 KIQVQKMYQVRIQPSIHQHNIPLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYED 1002
987 -ERLPSHSD--FLAELRKAPVNCSTAVCQRIQCDIPFFGICQEEFNATLKNLSFDWYIK 1043
1003 LERLPAARPCPLGALFRCPV-----FRQELIVQVIGTLELVGEIE 1044
1044 TSHNHLIVSTABILFNDVSFTLLPQGGAFVRSQETKVEPFEVFNPLPLIVGSSVGGLL 1103
1045 AS-SMPSLCSLSISFNSKHFHYGNSNL-AQVMKVDVYVEKQMLYLYVLSGIGGULL 1102
1104 LLALITAAALKYLGFFKQYKDMMSSEG-GPPGAP 1136
1103 LLLLIFVLYKVGFFKRNILKERMEAGRGVNGIP 1136

RESULT 5
I56126

lymphocyte function-associated molecule-1-alpha - mouse
C:Species: Mus musculus (house mouse)
C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I56126
R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A:Reference number: I56126; MUID:91268576; PMID:2051027
A:Accession: I56126
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1163 <RES>
A:Cross-references: UNIPROT:P24063; GB:M60778; NID:G198785; PIDN:AAA39426.1; PID:G198785
C:Genetics:
A:Gene: LFA-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F:151-315/Domain: von Willebrand factor type A repeat homology <VWAL>

Query Match 26.0%; Score 1527.5; DB 2; Length 1163;
Best Local Similarity 34.1%; Pred. No. 9e-98;
Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;
QY 1 FNLDTENAMTFOENA-RGFQSVVOLQGSRVVVGAPQIIVAAQORGSILYQCDYSTGSCPE 59
Db 24 YNLDTPTQSLAQAGRHFGVQLIEDG-VVGAPGE---GDNVTGGLYHCRSTSEFCQP 79
QY 60 IRLQVPEAVNMSLGLSLAATTSPQLLACQPTVHQTCSENTYVYKGLCFLFGSNLRQOPQ 119
Db 80 VSLH-GSNHTSKYLGWTLATDAAGSLLACDPGLSRCDQNTYLSGLCYLFPQSLGEGPML 138
QY 120 KFPALRGCCQEDSDIAFLIDGSGIIPHPRMKWVSTVMEOLKSKSLTFLSMQYSEE 179
Db 139 QNRPAYQECMKGVLDVLFDFGQSOLDRKDFEKILEFMKOWMRKLSNTSYQFAAVQSTFD 198
QY 180 FRIHFTPEKEF-QNNPNRSLIKPITQLGRTHATATGLKRVVRELENIINGARKNAF 238
Db 199 CRTFTFLDVVKQNKPDVLLGVSQPMFLTNTFRVYVAVHVFKEESGARPDATKVLV 258
QY 239 LLTDGEKF--GDPLGYEDVPIELDREGVIRVVGDAFSEKSRQBELNTVASKPPRDHV 296
Db 259 IITDGEASDKGNISAAHD-----ITRYIIGIKHFVSQKQKTLHLIPASEPVEFV 309
QY 297 FQINNFEALKTIONLREKIFALETGTGSSSPHEHMSQEGSAAITNSGPLLSTVGSY 356
Db 310 KILDTPEKLDLTELQKLYIVIEGTSKQDLSFNMLSSSGISADLSKGHAVVAVGAK 369
QY 357 DWAGGVF-LVTSKEKSTFTINMTRVDSMDNDAYLGYAAA-IILNRVQSLVIGAPRQ 414
Db 370 DWAGGFLLDLEDLQAGATFVQEBPLTSDVRGGYLGTVVAMTSSRSRPLLAAGAPRYQHV 429
QY 415 LVAMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLLVIGAPHYETRG 472
Db 430 QVLLFOAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGEAEALLIGAPLFGEQ 489
QY 473 GQSVSCPLPRGQARWQDAVLYGEOGQPMGRFGAALTIVLGDVNGDKLTDVAIGAP 532
Db 490 GRVFTY---QRRQLGFVEVSELDQDPGLRFGAALTALTDINGDGLVDVAVGAPLEB- 545
QY 533 NRGAVYLPHGTSGSGISPSHSQRJAGSKLSPLQYFGQSGGQDLTMDGLVDLTVGAOG 592
Db 546 -QGVYIFNGKPG-GLSPQSPQRIQAGVFPGRIRMFGRSHGVKDLBGDGLADVAVVGAEG 603
QY 593 HVLLRSQPLRVKAINMEFNPREVARNVFECDQV--KGKAGEVRVCLHVQKSTRDLR 651
Db 604 RVVWLSRPVVDVVTLSFSPETIPVHEVECSYSAREKQKHGKVLKACPRILTPQ--F 661
QY 652 EQIQSVTVYDLALDSGRPHSRVAVENETKSTRQTVLGTQTCETLKLQLPNCIEDPV 711
Db 662 QGRLLANLSYTLQDGHMRSGRLFPDGSHELSGNITSITP-DKSCLDHFHFPIQIDLI 720
QY 712 SPIVLRNFSLV---GTPLSAFGN-LRPVLAEADAQRLFTALFPEKNCGNNDNICODLSI 767
Db 721 SPINVSINLSLWEEGTFRDQRAQGDIPILRPSLHSETWEI-TKEIPFEKNGCKKCEANLTL 779

```

768 TFSFMSLDCLVGGP-----REFNVTVVRNDEGDSYRTQVTFPPPLDLISYRKVSTLQN 821
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
780 SSPARS-----GFLRMSSASLAVETWTLNSGEDAYWVRDLDFPRGLSFRKVEMLQ- 831
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
822 QRSORSWRLACESASSTEVSGAL-KSTSCSINHPPIFENSEVTFNITFDVDSKASLGNKL 880
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
832 --PHSRPVSCEEL--TEGSSLTKTKCNVSSPIFKAGOEVSQVWENILNLSWEDFV 887
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
881 LLKANVTSEN-NMPTNKTEFQELPVKYAYVMVTVTSHGVSTKYKYNFTASENTSRVMQHQ 939
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
888 ELNGTVCHENENSLQEDNSAATHIPVLPVNIILTKQENSTLYISFTPKGPKTQQQVHV 947
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
940 YQVSNLQORSPLISLVLVPVRLNQTVIWRPQ-----VTFSENLS--TCHTKE-RLP 990
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
948 YQV-----RIQPSAYDHNMPT-LEALVGPVRPHSEDLITYTWSVQTDPLVTVCHSEDLKRP 1001
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
991 SHSDFELAEKAPVNVCSIAVCQIQCDIPFGIOEENATLKNLSFPDWIKTSHNHL 1050
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1002 SSE--AEQPCPLPGV-----QPRCPFV--RWEILIQVTGVTELSKEIKAS-STLS 1046
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1051 IVSTAELLFNDVSFTLLPGQAFVRSQTEKTVPEPEVENPILPLIVGSSVGGLLALLITA 1110
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1047 LCSSLVSFNSKPHLYGSKA-SEAQVLVKVDLHERKEMLHVYVLSGIGGLVLLFLIFL 1105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1111 ALYKLGFFKQYKQMM-SEGPPGAEP 1136
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1106 ALYKVGFFKRLKEXMEADGGVPNGSP 1132
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 6
A53213
integrin alpha-E chain - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C:Accession: A53213
R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit.
A:Reference number: A53213; PMID:94164962; PMID:8119947
A:Accession: A53213
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: UNIPROT:P38570; GB:L25851; NID:9457244; PID:9457245
C:Genetics:
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.4%; Score 1140; DB 2; Length 1179;
Best Local Similarity 28.9%; Pred. No. 8.6e-71;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATSPPPOLLACGTVHQ 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 GPLHRCSLVQDEILCHPVEHVPIPKGRHGVTVVRSHGVLCI-----QVLVRRP--HS 117
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 96 TCSNTYVKGCLFLGSLNRQQPQ-----
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 LSSELT---GTCSLGGLDLPQQAQNFLENLLDPARVDVTDGCYNKKGEGDDVNTA 174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 KFPALRGCPQED-----SDIAFLIDGSGIIPFRRKKEWSTVMEQI-- 164
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 175 RQRALKEEEDKEEEDDEEAEAGTEIALIDGSGIDPPDFQRAKDFISNMWRNFE 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 165 KKSRTLSLMQYSBEFRIHFTTFKEFQNNPNRSLIKITQLLGHTHATGLRKVVRLEFN 224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 235 KCFCNFALVQYGGVIQTEFLDRDSQVMASLARQVNIITGVSVTKTASAMQHVLDISFT 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 ITNGARKNAFKILLFLITDGEFGDPLGYEDVIPELDREGVIRYVIGVGDAPRSEKSRQEL 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

295 SSHGSRKASKVMVLTLDGGIFEDPLNLTTVINSPKMQGVVERFAIVGVEFKSARTAREL 354
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 NTVASPPRPHVQINNFEALKTIQNLREKIFAIEGTQTGSSSSSEHEMSQEGFSAAIT 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
355 NLIASPDETHAPKVTNIMALDGLLSKLRNIISMEGT---VGDALHYQLAQIGFSAQIL 411
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
345 SNGP-LLSTVGSVDWAGGVFLY--TSKEKSTFINNTRVDSMDNA-----YLGYAAAIILRN 398
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
412 DERQVLLGAVGDFWMSGALLYDTRSRGRFLNQTAATAAAADAAQAQSYLYGYAVAVLHKT 471
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
399 RVQSLVLGAPRYQHI GLVAMFR-ONTGMWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
472 CSLSYAGAPQYKHG--AVFELQKEGREASFUPVLEGEQMGSTFGSELCPVDIDMGST 529
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
457 DLVLIGAPHYYEQTRGGQVSVCPPLRQGRARWCQDAVLYGEGQGPWGRFGAALTVLGDNV 516
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
530 DFLVLAAPFVHVHGEGRVYVRLSE-QDGSFSLARILSGHPGFTNARFGFMAAAGDLS 588
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
517 GDKLTDVAIGAP-----GEEDNR--GAVYLFHGTSGSGLSPSHSQRISAGLSPLRQYFGQ 570
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
589 QDKLTDVAIGAPLEGFGAGDGCASFSGSVIYNG-HWDGLSASFQIRASTVAPGLQYFGM 647
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
571 SLSGGQDLTMDGLVDLTGVAQHVLLLRSQPLVAKAIMEFNPREVARNVFECNDQVVKG 630
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
648 SMAGGFDISGDGLADITVTILGQAVVFRSRVRLKVMATFESALP-----LGF 697
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
631 KEAGEVRVCLHVQKSTRDLREGQIQSVTVTVLDALDGRPHSRVAVNFETKNSRTOVL 690
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
698 NGVNVRLCFEI-SSVTTASEGLREALLNLTLDVGVKQRRRLQCSVDVRSCLGCLREWS 756
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
691 GLTQTCETLKLQLPN-----CIEDPVSPVILRNFLNLSLGTPLSAFGLNRPVLAEDARLFT 746
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
757 SSSQLCEDL-LLMPTGEGELCEDCFSNASVSYQL-QTPEGQTDHPQILDRTYTEPPAI 814
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
747 ALFPEKKGNDNIQDDLSITFSMSLDCLVVGGRPRENVTVTVVRNDEGDSYRTQVTF 806
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
815 FOLPYEKAKCKKLFCVAELQLA-TTVSQQLVVGTLKELTILNLTNSGEDSYMTSMALN 873
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
807 FPLDLSYRKVSTLQNRQSRWRLACESASTEVSGALKSTSCSINHPIFPENSEVTFNI 866
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
874 YPRNIQ-----LKKMQKPPSPNIQDDDPQPV---ASVLMNCRIGHVPL-KRSSAHVS 923
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
867 TFDVDSKASLGNKLLKANVTSENN-----MPTNKTEFQ---LELPVKAVVMVYVTHGV 919
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
924 VQOLENAPNRTADITVTVTNSNRRSLANETHTLQFRHGFVAVLSKPSIMVNTVGGQL 983
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
920 S---TKYLNTASENTSRVMQHQVSNLQORSPLISLVLVPVRLNQTVIWRPQVTFSE 977
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
984 SHHKEFLPHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLRTQ 1028
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
978 NLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQIQCDIPFGIQEEFNATLKNLS 1037
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1029 ASTVCTWSEERACAYSS-VQHVEEHSVSCVIA-----SDKENVTVAEIS 1073
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1038 FQWYIKTSHNHLIVST-----AEILFNDVSFTLLPGQAFVRSQTEYKVEFFVNPCL 1091
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1074 WD-----HSEELLKDVTELQILGEISFNKSLGLENAENH--RTKITVTVFLKDEKYHSL 1125
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1092 PLTVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSE 1128
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1126 PIIKSGVGGLLLVILVILFKCGFFKRYQIQLNLE 1162
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7

A45226
integrin alpha-1 chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45226
E:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 269, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit

A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1151 <BRI>

A:Cross-references: UNIPROT:P56199

A:Experimental source: hepatoblastoma cell line HepG2

A:Note: sequence extracted from NCBI backbone (NCBIP:124326)

F:142-317/Domain: von Willebrand factor type A repeat homology <WAL>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;
Best Local Similarity 27.4%; Pred. No. 1.5e-67;
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

```
QY 1 FNLDTENAMTFQENARG-FGQSVVOL---OQSRVVVVGAPQEIIVAAQNRSLQCCDYSTGS 56
Db 1 FNVVDKNSMTSGPVEDMEGYIVQYENEEGKWKVLIGSLVGPQKRNRTGDVYKCPVGRGE 60
QY 57 CEP-IRLOPVBEA-----VNMSLGLSLAATSPQOLLACGPTVHQTCSNTYVKGIL 106
Db 61 SLPCVKLDLPVNTSIPNVTEVKENMTFGSTL-VTPNPGFLACGLYAYRCGHLHYTGI 119
QY 107 CFLFGSNLRQOPQKPEALRQCPQEDSDIAFLIDGSGSIIPHDFPRMKEWSTVMEQLKK 166
Db 120 CSDVSPPTQVNSIAP--VQBCSTQ-LDIVIVLDGNSIYPWDS-----VTAFLNDLLK 170
QY 167 -----SKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPTITOLLGR--THATGLRKVV 219
Db 171 RMDIGPKQTQGVIVQYGENVTHEFNLNKYSSTEEVLAAKLVQGGQRTMTALGTDTPAR 230
QY 220 RELFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRYVIGVGDAPF--- 276
Db 231 KEAFTEARGARRGVKVMVITDGBSH-DNHRLLKVIQDCDENIQRFSIALGSYNRGN 289
QY 277 --SEKSRQELMTVASKPRDRHVFQINNPEAKTIONQRKIFAEGTQTGSSSFHEHM 334
Db 290 LSTKFEVEIKSIASEPTKEHFNFNVDLAVTVIKTLGERIFALEATADQSAASFEMEM 349
QY 335 SQEGFSAITNGPLLSLVGSYDAGVFLYTSKE-----KSTP-INMTRVDSMDNDAYL 388
Db 350 SQTGSAHYSDQWMLGAVGYDNGVTVMQKASQIIIPRNTTINVESTKNEFL-ASYL 408
QY 389 GYAAAIILNRVQSL-VLGAIFYOHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCS 447
Db 409 GYTNSATASGDVLYIAGQPRYNHTGVLIYRMEDGNKIKLOTLSGEIQSGYSGILTT 468
QY 448 VDSDNSGTDVLVIGAPHY-----YEQTR-GGQVSVCPPLPRGQARWQCD 491
Db 469 TDIDKDSNTDILLVGAPMYMGTEKEBQGVVYVYALNQTREYQMSLEPIKQTCSSRQHN 528
QY 492 AVLYGEOQPMWG-RFGAALTVLGVDWGDKLTDAVAGPGEEDNRGAVLYLHGTSGSGISP 550
Db 529 SCTTENKNEPCGARFGTAIAVKDNLNDGNDIVIGAPLEDDHGGAVYIYHG-SGKTIRK 587
QY 551 SHSQRISAGLSPLQVFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSOPVLEVKAIM 610
Db 588 EYAQRIPIGGDGKTLKFFQSGIHEMDLNGDGLTDVTIGLGAALFWSRDVAVVKVTMN 647
QY 611 FNPREVARNVPCNDQVVKKEAG--EVRCVLHVQ-KSTRDRLEGOIQSVVTVDLALDS 667
Db 648 FEPKNVIQKNCH---MEGETVCINATVCFEYKLSKEDTIYEADLQ-----YRVTLDS 700
QY 668 GRHSRAVFNET-----KNSTRQTVGLTQTCETLKLQLPNCIEDPVSVLRNLFNS 721
Db 701 LRQISRSFFSGTQERKQVORITVKSEC-----TKHSFYMLDHDFODSVR---ITLDEN 752
QY 722 LVGTPLSAPGNLRVLAEDAQRFLTALFPFEKNGCNDNIQDDISITTFSEWSDCLVWVG 781
Db 753 LT-DPENG-----PVLDDSLPNSVHEYIPPAKDCGNKEKICISLSLHVATEKDLIVRS 806
QY 782 PRE-ENVTVTRVDGDSYQTVPFFPLDSLVRKVTSTLQNRQSRQSRWLACSSASTEV 840
Db 807 QNDKFNVSLLVTKOSAYNTRTIVHYSPLNLFVSGIEAIQKD-----SCSEN----- 853
```

RESULT 8

145914

Integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004

C:Accession: 145914

R:Kamata, T.; Fuzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integrin

A:Reference number: A54402; MUID:94193647; PMID:7511592

A:Accession: 145914

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <KAM>

A:Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g433691

F:161-336/Domain: von Willebrand factor type A repeat homology <WVA>

Query Match 18.2%; Score 1071; DB 2; Length 1170;
Best Local Similarity 27.4%; Pred. No. 5.5e-66;
Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

```
QY 1 FNLDTENAMTFQ-ENARGFGQSVVOL---OQSRVVVVGAPQEIIVAAQNRSGSLYQC--DYST 54
Db 19 YNVGLPKAKIPSGPSSQSGFGYAVQOFTNPKGNMLLVGSPWSPGPKNRMGDYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQOLLACGPTVHQTCSNTYVKGILC 107
Db 79 TTCEKLNQITSTMSNVTEKTNMSLGLTLTRNVGTGGFLTCGFLPAQCGSQYTTTGV 138
QY 108 FLFGSNLRQOPQKPEALRQCPQEDSDIAFLIDGSGSIIPHDFPRMKEWSTVMEQLK-- 165
Db 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVVCDENSIYEPD--AVKNFLEKFEVQGLDIG 194
QY 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPTITOLL-----GRTHATGLRKVVRE 221
Db 195 PTQKQGLQYANNPRVFNLTNFKSKD---EMIKATQTFQYGGDLTNTFKAIQYARDT 251
QY 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVPELDRGVIRYVIGV-----GDAFR 276
Db 252 AYSTAAGRGATKVMVVVTDGESH-DGSKLKAVIDQCKNDILRFGIAGVLYNRLALD 310
QY 277 SEKSRQELMTVASKPRDRHVFQINNPEAKTIONQRKIFAEGTQTGSSSFHEHMSQ 336
Db 311 TKLKIKEIKAIASIPTRHFFNFVNSDEADLEKAGTTIGEIQIFSIEGTVOG-GDNFQMEMSQ 369
QY 337 EGFSAAIT--SNGPLLSLVGSYDAGVFLYTSKEASTFINMT--RVSDSMN-DAYLGYA 391
```

370 VGFSAEYSPQNNILMLGAVGAYDMSGTVVQKTPHGLHIFSKQAFQEIILQDRNHSYLGY 429
392 AAILRNVRQSLVLGAPRYOHIGLVAMPFRONTGMWESNANY-----KGTOIGAYFGASL 445
430 VASISTGNSVHFVAGAPRANTVQGVLYSVN-----ENGNTVTIQSORQDQIGSYFGSVL 484
446 CSVDVDSNGSTDLVLGAPHYYEQTR--GGQVSVCLPRGQARQCDVAVLYGEGQDPWG 503
485 CAVDVKNDTITDVLVAGPMYMDLKKBEGRVLETTIKG-ILNWH--QFLEGGNLENA 541
504 RFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVLFGHTSGSGISPSHSORIASG--KL 561
542 RFGSAIALSLINDMGFNDVIVGSPLEQNSGAVIYNGHEGM-IRLAYSOKILGSDRAF 600
562 SPRLOYGSGGGQDLTMDGLDVLTVGAQGHVLLRSQPVLRVKAIMEFNPREFVARNVF 621
601 SSHLQYFGRSLDGYDGLNGSDITDVSAGFAFGVQLWSQSIADVSADSFTEPKKI--TLL 658
622 ECNDQVVKKEAGEVRVCLHVQKSTRDLREGIQSVVTDIALD-----SGRPHSRVFN 677
659 NKNABE-----KLKLCF-----SAKFRPTQNNQVAIVNITIDEDQFSSRVISRGLEK 707
678 ETKNSTRQTVLGLTQTCB--TLKILQPLNCIEDPVSPIVLRNLSL--VGTPLSAFGNL 733
708 ENNERCLQTMIVSOAQRCEYIIHQEBS--DIISPLNLCWNISLENPGT-----756
734 RPLVLEDAQRLFTALFPFKKNGCNDNI CODDLSITP-----SPMSLDCLVVGPRFENVTV 789
757 NPALAEYSETVTVFSIPFHKDGDGVCISDLVAVQQLPATQQQPFVIVSNQNKLTFSV 816
790 TVRNDEGDSYRTQVTFPFFPLDLSYRKVSTLQNRQSRWSRLACESASST-EVSGALKSTS 848
817 QLNKKESAYNTEIIVDFSENLF-----ASWMPVDGTEVTCQIASSQKSVT 864
849 CSINHIPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTETQLELPVKY 908
865 CNVGYPAKSKQQTFTINFDFNLQ-NLQNOASISPRALSESQEEENMADSNVNLKLSILY 923
909 AVYVMVTSIGVSTKYLNFTASENTSRVMOHQVSNLQOR-----SLPISLVFLV 958
924 DAEIHIT-RSTNINFEVSLDGNVSVV-HSFE--DIGPKFTFSIKVTGTSVVPVSNVMA---976
959 PVRLNCTVWRDPQVTFSEN--LSSTCKE-----RLPSSHDFLAE- 998
977 -----SVIHIPIQYTKDNKPLMLTGCVHTDQAGDISCAEINPLKIGQTSSSVSFKSEN 1030
999 LRKAPVNGSIAYCQRIQCDIPFGIOEBFNATLKNLGFDPYIKTSHNHLILIVTAEI- 1057
1031 FRHIKELNCKTASCNSIMCWLRLQVKGEYFLNVSTRIWNGTFAASTFTQVQLTAAAEID 1090
1058 LFNDSVFTL-----LPGQAFVRSQTETKVEPFE-VPNPLPLIVGSSVGGILLALLITA 1110
1091 TYNPQIYVIBENTVTP-----LTIMKPEHKEVPTGVIVGSVIAGILLALLALVA 1140
1111 ALYKLGFFKRYQKDM 1125
1141 ILWKLGFFKRYKEM 1155

RESULT 9

S44142

VLA-2 protein homolog - mouse

C;Species: Mus musculus (house mouse)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S44142

R;Edelman, J.M.; Chan, B.M.; Uniya, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze

submitted to the EMBL Data Library, January 1994

A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not

A;Reference number: S44142

A;Accession: S44142

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:G473098; PID:CAA82877.1; PID:G473
P;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;

Best Local Similarity 27.8%; Pred. No. 5.3e-65;
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;

QY 1 FNLDTEAMTFQ--ENARFGQGVVQL-----QGSRRVVVGAPOEIVAAANQSGSIYQC--DYST 54
DB 27 YNVGLPGAKIIFSGSSSEQFGYSVQQLTNPQGNWLLVGSFSGFPPENRMDGVVYKPCVDLPT 86
QY 55 GSCSPIRLO-----VPVEAVNMSLGLSLAATTSPPOLLACGPPVHOTCSENTVVKGLC 107
DB 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPTGGFLTCGPLWAHQCGNQYVATGIC 146
QY 108 FLFGSNLRFQOQ--KPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMVSTVMBSQL 164
DB 147 ----SDVSPDFQFLTSFSPVAQCDLSL--VDVVVCDENSIYP--WEAKNKLFLVKFVTGL 199
QY 165 K--KSKTLFSLMQXSEEFRIHFTFEKQNNPNRSLIKPIITQLLG-RTHATGLKRVRE 221
DB 200 DIGPKKTOVALIQYANEPRIIFNLNDFETKEDMVQATSETROHGGDLTNTFRAIFARDY 259
QY 222 LFNITNGARKNAFKILFLLTGHEKGDPLGVEDVIPELDREGVIRYVIGV-----GDAFR 276
DB 260 AVSQTSGRPGATKVMVVVTDGESH-DGSKLKTIVQQCNDDEILRFGIAVLGYLNRLD 318
QY 277 SEKSQELNTVASKPRDHFVFOINNFEALKTIONLRKIPAIETGQTGSSSFHEMSQ 336
DB 319 TKMLKEIKATIASTTERFFNVDEBAALLEKAGTLGQIFSGIEGTVGQ-GDNFQWEMAQ 377
QY 337 EGFS--AITSNGPLLSLVGSYDAGGVFLYTSKSTFINMT--RVSDMN-DAYLGYA 391
DB 378 VGFSAADYAPQNDILMLGAVGAFDWSGTLVQETSHKPPVPPKQAFQVLDNRHSSPLGYS 437
QY 392 AAILLRNVQSLVLGAPRYOHIGLVAMPFRONTGMWESNANY-----KGTOIGAYFGASLCS 447
DB 438 VAAISTEDGVHFVAGAPRANTVQGVLYSVNK--OGNVTVIQSHRGDQIGSYFGSVLCS 494
QY 448 VVDVDSNGSTDLVLGAPHYYEQTR--GGQVSVCLPRGQARQCDVAVLYGEGQDPWGFR 505
DB 495 VVDVDDKTTIDVLVAGAPTYMNDLKEEGKVLYFTTKILNQHQ---FLEGEGTGNARF 551
QY 506 GAALTVLGVNGDKLTDVAIGAPGEDNRGAVLFGHTSGSGISPSHSORIASKLSR- 564
DB 552 GSAIAALSINMDGENDVIVGSPVENENSGAVIYNGHEGT-IRTKYSOKILGNSGAFRR 610
QY 565 -LQYFGQSLSGQDLTMDGLDVLTVGAQGHVLLRSQPVLRVKAIMEENPREVARNVFEC 623
DB 611 HLQFFGRSLDGYDGLNGSDITDVSIGALGVQLWSQSIADVAIEALFTP-----660
QY 624 NDQVVKKEAGEVRVCLHVQKSTRDLREGIQSVVTDIALD-----SGRPHSRVFNET 679
DB 661 -DKITLLNKDAKITLKLCFRPFEPAGQNNQV--AILFNMTLDADGHSRVTSGRVFREN 717
QY 680 KNSTRQTVLGLTQTCET--LKLQPLNCIEDPVSPIVLRNLSLIVGTPLSAFNLRLPVL 737
DB 718 SERFLQKNVMVNEVQKSEHHISIQKPS--DVVNPLDLRVDISLENPGTS-----PAL 768
QY 738 AEDAQRLFTALFPFEKNCNDNICODDLSI-----TFSPMSLDCLVVGPRFENVTV 788
DB 769 EAYSETVKVFSIPFYKECGSDGICISDLILDVQQLPAIQQSF-----IVSNQKRLTFS 823
QY 789 VTVRNDEGDSYRTQVTFPFFPLDLSYRKVSTLQNRQSRWSRLACESASST-EVSGALKST 847
DB 824 VILKNGESAYNTVILAESENLF-----ASFSPMDVGTGTEVTCVSGSSQKSV 871
QY 848 SCSSINHIPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET--FOLELP 905
DB 872 TCDVGYPAKSKQQTFTINFDFNLQ-NLQNOAIAINFQAFSESQ--ETNKADNSVSTIIP 928
QY 906 VKYAVVMVTSIGVSTKYLNFTASENTSRVMOHQVSNLQOR-----SLPISLV 955

Db 929 LLYDAELHLT-RSTWTFNFEISSDENAPSVIK---SVEDIGPKFIPSLKVTAGSPVSWA 984
 QY 956 FLV-----PVLNQTVIWDPPQVTF-SENLS 980
 Db 985 LVTHIHPOYTKENPLLYLTGTOQAGDISCTAEINPLKLPHTA---PSVSFKNEFR 1040
 QY 981 STCHTKERLPSPDSFLAELRKAPVNVCSIAVQRIQCDIPFFGIOBEFNATILKGNLSFDW 1040
 Db 1041 ---HTKE-----LDCRTTSCNITCWLKDLHMAKAEYFINVTRVWART 1080
 QY 1041 YIKTSHNHLIVSTAEILLFNDVSFTLLPQGGAFVRSQTEKVEPEVFNPLPIVSSVG 1100
 Db 1081 FAASTFQVQTAAAEIDTHNPQLEFVIEENAVTIPILMIMKPTKEAEVPT--GVILGSIITA 1138
 QY 1101 GILLHALLITAALYKLGFFKQVYKDM 1125
 Db 1139 GILLHALLANTAGLWKLGFQVYKDM 1163

RESULT 10
 A35854
 integrin alpha-1 chain precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
 C/Accession: A35854; S11243
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
 A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
 A:Reference number: A35854; MUID:90338125; PMID:2380249
 A/Accession: A35854
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1180 <IGN>
 A/Cross-references: UNIPROT:P18614; GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
 F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 17.9%; Score 1055; DB 2; Length 1180;
 Best Local Similarity 27.4%; Pred. No. 7,3e-65;
 Matches 346; Conservative 196; Mismatches 475; Indels 246; Gaps 48;

QY 1 FNLDTENAMTQENARG-FGQSVVQL---QGSVVVVGAPQBIVAANQSGSIYQCDYSTGS 56
 Db 29 FNVDFKNSMSPSGPDMFGYTVQYENECKWVLIGSLGSLVPGQPKARTGVYKCPVGRER 89
 QY 57 CEP-IRLQVPEA-----VMSJGLSLAATTSPPOLLACGPTVHOTCSENYVYKGL 106
 Db 89 AMPCVKLDLPVNTSIPNVTEIKENMTFGSTL-VTNPNNGGFLACGFLYAYRCGHLHYTTGI 147
 QY 107 CFLFGSNLRQOPKFPPEALRCPOEDSDIAFLIDGSGSIIPHDFRMKEWSTVM---EQ 163
 Db 148 CSDVSPTFQVNSFAP--VQECSTQ-LDIVIVLDGNSIYP-----WESVIAFLNDL 196
 QY 164 LKK-----SKTLFSLMOYSEFRIHFTFKGFQNNPNRSLIKPITQLLG-RHTPATGLRK 217
 Db 197 LKRMIDGFKQTGVQIVQYGENVTVEFNLNYSSTEEVLVAANKIGRQGLQTMALGIDT 256
 QY 218 VVRELENTINGARKNAFKILFLTDGKFGDPLGYEDVIELDREGVIRVIVGVGDAPR- 276
 Db 257 ARKEAFTEARGARKVKVMVIVTDGESH-DNVELKQVIOQCEDENIQRFSAIILGHYNR 315
 QY 277 ----SEKSRQELNTVASKPRDHVFQNNFPEALKTIQNLREKIFATEGQTQSSSFEH 332
 Db 316 GNLSTKRFVEIKSIASEPTKEHFFNVSDDELALVTIVKALGERIFALEATDQSAASFEM 375
 QY 333 EMSQEGFSAITSNGLPSTVGSYDWDAGGVFLYTSKEKSTFINMT--RVDSMDND---AY 387
 Db 376 EMSQTGFSAHYSQDWMLGAVGAYDNGVTVMQANQVIPHNTTFTQTEAKMNEPLASY 435
 QY 388 LGYAAAILLRVQSLVGLAPRVQHIGLVAMFRONTGWESNANVKQTQTGAYFGASLCS 447
 Db 436 LGYTVNSATIPGDVLYTAGQFRYNTHTQVVIYKMDNGNINILQTLGGEQIGSYFGSVLTT 495

QY 448 VVDVDSNGSTDLVLICAPHY-----YEQTR-GQVSVVCPLPQRQARWQCD 491
 Db 496 IDIDKDSYTDLLLVGAPMYMGTEKEEQGVVYAVVQTRFEYQMSLEPIRQTCSSLKDN 555
 QY 492 AVLYGEGQOPWG-REGAALTVLGDYNGDKLTDVAIGAPGEEDNRGAVILFPGTSGSISF 550
 Db 556 SCTKENKNEPCGARGETAIAAVKDLVNGFNDVIGAPLEDDHAGAVIYIHG-SGKTIRE 614
 QY 551 SHSQRIAGSKLSPRIQYFGQSLSGQDLTMGDLVLTGGAQGHVLLLRSQPVLVRKAIM 610
 Db 615 AVAQRIPSGGDKTKLFFGQSHGEMDLNGDLTDVTIGLGAALFWARDVAVVKTWN 674
 QY 611 ENPRVARNVPECNDQVVKGEAG--EVRVCLHVQ-KSTRDRLRGQTSVVVYDLDLSD 667
 Db 675 FEPKNVNTQKXNCR---VEGKETVCINATMCFHVKLKSKEDSIYEDLIQ---YRVTLDS 727
 QY 668 GRPHSRVAFNET-----KNSTRQTOVLGLTQTCETLKLQLPNCI-----EDPVS 712
 Db 728 LRQISRSPFGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQD 770
 QY 713 PIVLRNLNPLVGTPLSAFNGLRPVLAEQAQLFTALFFPEKNCNGNDNICQDDLSITFSFM 772
 Db 771 SVRVTLDELNT-DEPENG-----PVLDDALPNSVHEHIFPAKDCGNKERCISDLTLNVSTT 824
 QY 773 SLDCLVYGGPRE-FNVTVTVNDGEDSYRTQVTFPPFDLDLSYRKVSTIQNORSQSWRLA 831
 Db 825 EKSLLIVKSKHQDKFNVLTVKNGDSAYNTRTVVQHSNPLIFSGIEETQKD-----S 876
 QY 832 CESASSTEVSGALKSTCSINHPIFFENSEVTEINTFDVDSKASLGNKLL-LKANVTSEN 890
 Db 877 CESN-----QNTICRVGYPLRAGETVTFKLIIFQFNTSHLSENAIHLSTSDSEE 927
 QY 891 NMPTNKTEFOLELPVKYAV-----YMWVTSHGVT-----KVLNFTASENTSRVMOHQ 939
 Db 928 PLESINDNEVNISSIPVYEGVLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVP 985
 QY 940 YQVSNLQORSILP---ISLVP-----LVFVRLNQTVIWD-----RP----- 971
 Db 986 YTIRKRGHPMPQLSISFPNLTAGDGPVLYPIG-----WSSSDNVNCRPSRLEDPFG 1039
 QY 972 ----QVTF-----ENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQ 1016
 Db 1040 INSKKMTISKSEVLKRGTTQDCSTC-----GVATITCSLLPSDLSQ 1082
 QY 1017 CDI-----PFFGQIEEF---NATLKNLSFDWYIKTSHNHLIVSTAEILLFNDVSFTLL 1067
 Db 1083 VNVSLLLWKPTF-IRAHFSSILNLTGELK-----SENSLTLSSN----- 1123
 QY 1068 PGQGAFVRSQTEKVEFEVFNPLPL--IVGSSVGGLLLLALITALYKLGFFKQVYKDM 1125
 Db 1124 -----RKRELAIQISKDGLPGRVPLVILLSAFAGLLMLLILALWKIGFFKRLPKKK 1177
 QY 1126 MSE 1128
 Db 1178 MEK 1180

RESULT 11

A33998

integrin alpha-2 chain precursor - human

N/Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 cl

C/Species: Homo sapiens (man)

C/Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004

C/Accession: A33998; B56793; A53117

R/Takada, Y.; Hemler, M.E.

J. Cell Biol. 109, 397-407, 1989

A/Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)

A/Reference number: A33998; MUID:89308879; PMID:2545729

A/Accession: A33998

A/Molecule type: mRNA

A/Residues: 1-1181 <TAK>

A/Cross-references: UNIPROT:P17301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907

A/Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803

R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GP1a, GP1c, GP1c*, GP1a and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
A;Zutterer, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'v', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-1133/Domain: extracellular #status predicted <EXT>
F:172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F:1134-1154/Domain: transmembrane #status predicted <TM>
F:1155-1181/Domain: intracellular #status predicted <CY>
F:1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;
Best Local Similarity 26.7%; Pred. No. 8.6e-65;
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFQ-ENARFGQSVVQL---QGRVWVGAPCEIVAAHQRSGLYQC---DYST 54
DB 30 YNVLPEAKIFSPSPSEQFGYVQVQFIPKGNWLLVGPSPFPENRMGDVYKCPVDLST 89
QY 55 GSCEPIRLQ-----VPFEVNMVSLGLSLAATSPQLLACGPTVHQTCSNTVYKGLC 107
DB 90 ATCEKLNQTSIPNVTETKMTNLSGLILTRNMGTGGLTCGLPLMAQQCGNYTTCV 149
QY 108 FLFGSNLRQOPQKFPALRCQPEDSDIAFLDGSIIIPHDFRRMKWSTVMEQLK-- 165
DB 150 SDISPPF-QLSASFSPATQCPSL-IDVVVVCDESIIYPWD--AVKNFLEKFPVQGLDIG 205
QY 166 KSKTLFSLIMQYSEPRIFHTFFKEFONNPNRSLIKPIITLLG-RTHTATGLRKVRELFN 224
DB 206 PTKTQVGLIQYANNPRVFNLTNYKTEEMIVATSTQSYGGDLTNTFGAIOYARKYAS 265
QY 225 ITNGARKNAKILFLITDGEKFGDPLGYEDVPELDREGVIRVIGV----GDAFRSEK 279
DB 266 AASGGRSRTKVMVWVVDGESH-DGSMKAVIDQCHNDNLIRFGIAVLGNRNALDTKN 324
QY 280 SROELMTVASKPRDHVFNLFNFEALKTIONQLREKIFAEGTQTGSSSFEHMSQEGF 339
DB 325 LIKEIKAIASIPERFFNVSDAALLEKAGTIGELIFSEGTIVQVQ-GDNFQEMSQVGF 383
QY 340 SAAITNGP--LLSTVGSYDWAGVFLYTSKESKSTFINMT--RVDSDMN-DAYLGAAAI 394
DB 384 SADYSSQNDILMLGAVGAFGWSGTIVQKTSGLHILFPKQAFDQILODRNHSYLGYSVAA 443
QY 395 ILNRVQSLVLGAPRYOHIGVAMFRONTGWESNANV-----KGTOIGAYFGASLCSV 448
DB 444 ISTGESTHFVAGAPRANTQOVLVSYN-----ENGNTIVQAHRGDQIGSYFGVLCV 498
QY 449 DVDNSGTDVLVIGAPHYEQTR--GGQVSVCPPLRCORARQCQDAVLVGEQGPWRFG 506
DB 499 DVDKDTITDVLVIGAPYMSDLKKEGRVYLFTIKKILGQHQ---FLEGPEGIENTRFG 555
QY 507 AALTVDGVNGDKLITVAIGAPGEDNRGAVLPHGTSGSGISPSHSORTAGS--KLSPR 564
DB 556 SAIAALSDINMGDFNDIVGSPLENQNSGAVIYNHGQGT-IRTKYSQKILGSDGAPRSH 614

565 LQYFGQSLSGQDLTMDGLVLTVAQGHVLLLSQPVLRVKAIMEFNPVREVARNVFCN 624
DB 615 LQYFGRLSDGYGLDNGSDITDVGAFQVQLWSQSIADVAIBASFTEPKI--TLVKNK 672
QY 625 DOVVGKEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALD-----SGRPHSAVFNETK 680
DB 673 AQII-----LKLCF-----SAKFRPTKONNQVAIVYNTITLADAGFSRVRTSRGLFKENN 721
QY 681 NSTRQTOVLGLTQFC--ETLKLQLPNCIEPVPSPVLRNLNFSILVGTPLSAGNLRPVLA 738
DB 722 ERCLOKNNVWQAQSCPHIYIQEPS---DVNSLDRVDSLENPGTS-----PALE 772
QY 739 EDAQLFTALPFFKXNCGNDNICQDLSITF---SFMSLDCLVVGGRPFNVTVTRND 794
DB 773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQOEOPFIVSNQNKRLTFSVTLKNK 832
QY 795 GEDSVRTQVTFPFLDLSYRKVSTLQORSQSWLACESAST-EVSGALKASTCSINH 853
DB 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
QY 854 PIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFOLELPVKYAVVMV 913
DB 881 PALKREQVTFINDFNLQ-NLQNASLSPQALSESQEEKADNLVNLKIPLLYDAEI- 938
QY 914 VTSHGVSTKYLNFATASENTSRVMQHQYQVSNLQGR-----SLPSLVFLV----- 958
DB 939 ---HLTRSTNINFEISSDGNVPSIVHSPEVGPKFIFSLKVTGTSVPVSMATVITHIQ 995
QY 959 -----PVELNQTIVWDRPQVTF-SENLSSTCHTKER 988
DB 996 YTKENPLMYLFGVQTDKAGDISCNADINPLKIGQT-----SSSVSPKSENFR---HTKE- 1047
QY 989 LPSHSDFLAELRKAIPWNCISIAVCORICDIPFFGQIEEFNATLKNLSFDWIKYIKTSHH 1048
DB 1048 -----LNCRTASCNVTCWLKDXVHKGEYFVNVTIRIWNKGFASSTFQT 1091
QY 1049 LLIVSTAETI-LFNDVSFTLLPQOGAFVRSQRTKVEPFPVNP-----LP--LIVGSSVG 1100
DB 1092 VQLTAAAEINTVNPFIYVI-----EDNTVTIPLMIMKPDKAEVPTGTIGSIIA 1141
QY 1101 GILLALALITAAALYKLGFFKRYKDM 1125
DB 1142 GILLALLVAILLWKLGLFFKRYKDM 1166

RESULT 12
A41131
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N;Alternate names: integrin alpha-4
C;Species: Mus musculus (house mouse)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C;Accession: A41131; S16742
J;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.
R. Cell Biol. 115, 1149-1158, 1991
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A;Reference number: A41131; MUID:92064645; PMID:1840602
A;Accession: A41131
A;Status: preliminary
A;Molecule type: mRNA
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:951484; PIDN:CAA37316.1; PID:9514
C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;
Best Local Similarity 22.7%; Pred. No. 1.3e-37;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDTENAMTFQ-ENARFGQSVV-QLQGRS--VVGGAPOEIVAAV----QRGLYQCDY 52
DB 41 YNLDPENALLYQSPGTLFGYSVVLHSGSKRWLVIGAPTASWLSNLSVNVNPGAIYRCGI 100
QY 53 STG----SCEPIRLQVP-----VEAVNMSLGLSLAATSTP-PQLLAG---PTVHQ 95

Db 101 RKNPQTCQLQSGSPGCPCKGTCLEERDQWGLVTLRSQPGENGSIIVTCGRWKNIFY 160
QY 96 TCSENTYVKGCLFCLFGSNLRQOPQPFPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKX 155
Db 161 MKSDNKLPTGICYVMPSDLRTELK----- 185
QY 156 WVSTVMEQLKSKTLFSLMOYSEEFRIHFTPEKQFNPNPRSLIKPITQLLGRTHATGL 215
Db 186 -----RMAPCYKDYT----- 195
QY 216 RRVVRELFNITNGARKNAFKILFLITDGEKFGDPLGYEDVPELDREGVIRYVIGWD 275
Db 196 -----RKFG----- 200
QY 276 RSEKSRQELNTVASKPRDHVFQINNPEALKTIQNLREKIPAIETGTGTGSSSPHEHMS 335
Db 201 -----NFAS-----C 205
QY 336 QEGFSAAITNSGPLLSTVGSYDWAGGVFLY--TSKEKSTFTINMTVRVDSMDNDAYLGXA-- 391
Db 206 QAGISSEFTQDLVWGAGSSVWTGTVFVYMITTNOYKAFVD--RQNVKFGSYLGYSVG 263
QY 392 AAILLRNVQSLVLGAPYQHIGLVAMFRQNTGMWESNAV---KGTQIGAYFGASLCS 447
Db 264 AGHFRSPHTTEVVGAPQHEQIGKAYIF---SIDENELIVYEMKGLSGYFGASVCA 319
QY 448 VDVDSNGSDVLVLGAPHYYQTRGGQVSVCPPLRGQRA--RWQCDVLYGQGPWGRFG 506
Db 320 VDLNADGSDLI--LVGAPMQSTIREGRVFVY--INSGMGVMMVEMERVLVSGDKYA--ARFG 376
QY 507 AALTVLGVNDGDKLTDLVAIGAPGEDNRGAVYVLTGTSSTGSGISPSHSORLAGSKLSPLO 566
Db 377 ESIANLGIDNDGFDIAIGAPQEDDLRGAVIYNGRV--DGISSTYSRIEQQOIKSLR 435
QY 567 YFGQSLSGQDLTMDGLVLDLVGA--QGHVLLRLSQPVRVLRKALMEFNPREVARNFECN 624
Db 436 MFGQISGQIDADNNGYVDVAVGAFQSDSAVLLRTPWVIVEASLS--HPESVNRKTKDCT 494
QY 625 DVVYKGEAGEVRVCLHVQKSTRDLRGQIQSVVTVDLALDSGR---PHSRVAF--NET 679
Db 495 -----ENGLPSVCMHLTLCFSYKGRVPGYIVLFYNVSLDVHRKAESPFRFFSNGT 547
QY 680 KNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVLRNLFSL-----VGTPLSAFGNLR 734
Db 548 SDVITGSIYSSSEKCKTHQAFNRKQVDRDILTFHVEATYHLGHVITKRNTEEFPPLO 607
QY 735 PVLAEDAQR--LFTALPFPEKNCNGNDNICODLSITFS-----FMSLDCLVVGPRFNV 787
Db 608 PILLQKKEKDVIRKMINFARFCAVEN--CSADLQVSAKVGFLLPYENKTYLAVGSMKTI 666
QY 788 TVTVRNDGEDSYRTQVTFPFPLDLSYKSVSTLQNRQSRWRLACESASSTEVSGALKST 847
Db 667 NVSLFNAGDDAYETTLNVQLTGLYFIKILDEEK-----QINCE-----VTESSGIVK-L 716
QY 848 SC SINHPIFPENSEVFNTITFDVDSKASLGNKLLKANVTSEN--NMPRNTKTEFQLELP 905
Db 717 ACSLGIYVDRLSRIDISFLIDVSSLSRAHEDLSIVHASCENEGELDQVRDNRVLTIP 776
QY 906 VKYAVVWVTSHGVI--STKYINFTASNTSRVMOHQ-----YQVSNLQGRSLP--ISLVFL 957
Db 777 LRYEV--MLTVHGLVNPSTSVYGSSENEPETCMAEKLNTLFHVINTGISWAPNVSKIM 834
QY 958 VPVRLNQTIVDRP--QVTFSENLSSTCHTKE-----RLPSHSDFLAELR 1000
Db 835 VP---NSFLPQDDKLFNVLDVQTTTQGCFFKHYGRECTFAQKQGIAGTLTDIVKFLSKTD 891
QY 1001 KAPVWNGSIA--VCQRIQCIDIPFGIOEENFATLKNLSFDWIKTSHNLLIVSTAEIL 1058
Db 892 KR--LLCYMKADQCHLDFLCN--FGKMEG-----KEASVHIQLEGRPSIL 933
QY 1059 FNDSVFTLLPGQAGFVRSOTETKVEPFEPVNP----- 1090

Db 934 EMDETSSLI-----KFEIKATAPPEPHPKVIELKNKDNVAHVFLGLHQRKRHF 983
QY 1091 -LPLIVGSSVGLLLALITAAIYKLGFFKQYKDMMS 1128
Db 984 TIIITISLLGLLIVLLIISCVMMKAGFFKQYKSILOE 1022
RESULT 13
S06046
integrin alpha-4 chain precursor - human
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain
C:Species: Homo sapiens (man)
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C:Accession: S06046; A39355; D28018
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.
EMBO J. 8, 1361-1368, 1989
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte
A:Reference number: S06046; MUID:89356603; PMID:2788572
A:Accession: S06046
A:Molecule type: mRNA
A:Residues: 1-1038 <TAK>
A:Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991
A:Title: Characterization of the alpha-4 integrin gene promoter.
A:Reference number: A39355; MUID:91239513; PMID:2034655
A:Accession: A39355
A:Molecule type: DNA
A:Residues: 1-93 <ROS>
A:Cross-references: GB:M62841
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A:Title: The very late antigen family of heterodimers is part of a superfamily of molec
A:Reference number: A94151; MUID:87204112; PMID:3033641
A:Accession: D28018
A:Molecule type: protein
A:Residues: 40-50, 'E', 52-53 <TA2>
C:Genetics:
A:Gene: GDB:ITGA4; CD49D
A:Cross-references: GDB:I28032; OMIM:192975
A:Map position: 2q31-2q32
C:Superfamily: integrin alpha-4 chain
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transm
F:1-39/Domain: signal sequence #status predicted <SIG>
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>
Query Match 10.98; Score 642; DB 2; Length 1038;
Best Local Similarity 23.88; Pred. No. 3.8e-36;
Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;
QY 1 FNLDTENAMTFQ-ENARGFGQSVV-QLQGSR--VVGAPQEIIVAAV-----QRGSLYQCDY 52
Db 40 YNVDTESALLYQGFHNTLFGYSVVLHSHGANRWLLVGAPTNLANASVINPAGIYRCRI 99
QY 53 STG---SCEPIRLQVPEAVNMVSLGLSLAATTFPQLLACGPTVHOTCSNTYVKGCLFL 109
Db 100 GKNPGQTCQLQGLSP-----NGEP-----KTCLEERDQWL--- 133
QY 110 FGSNLQQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMSVIVMEQLKSKT 169
Db 134 -GVTLRQPGE-----NGSIVTCGHR---W-----KN 156
QY 170 LFSLMQYSEEFRIHFTPEKQFNPNPRSLIKPITQLLGRTHATGLRKVVVRELFNITNGA 229
Db 157 IFVI----- 160
QY 230 RKNAPKIFLLTDGEKFGDPLGYEDVPELDREGVIRYVIGWDAPFRSEKSRQELNTVAS 289
Db 161 -KNEKN-----LPTGGCYGVF-----PDL-----RTELSKRIA 187
QY 290 KPRDHFVFOI--NNFEALKTIQNLREKIPAIETGTGTGSSSPHEHMSQGFSAITNSGP 348
Db 188 PCYQDIYKKGNFAS-----COAGISSFYTKDLI 217


```
QY 1040 WYIKTSNHLIIYSTAILELNDVSFTLLPGQGFVRSQSTETKVEP-----1084
Db 918 IY-----MLNT-EILKDDSSVIQ-----FMSRAKVVDPALRVEVIAHGNPEEV 962
QY 1085 ---FEVFN---PLPLIVG---SSVGGELLLALITAAALYKLGFFRQYKDDMM 1126
Db 963 TVVFEALHNLPRGYVVGWIIAISLLVIGLIFLLAVLLMKMGFFRRRYKEII 1015

RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 10.4%; Score 614.5; DB 2; Length 1041;
Best Local Similarity 24.8%; Pred. No. 3.1e-34;
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAAITNSGP--LLSTGVSYDWAGGVFLYTSKEKSTFINMTRVDS-----DMNDAYLG 389
Db 181 QAGFGLIFSDNSALVMGAPGSYYLQGIYQSLNRSV-VQATQESNTGYSPDNSYRG 239
QY 390 YAAAI--ILNRVQSLVGAPRYQHI-GLVAMFRQNTGMWESNANVKGTQICAYFGASLC 446
Db 240 YSLALGDFNGGVQDVYVGTFRASLMGLVAIFDQNLQFN--QVMGTQIVAYFGYSVT 296
QY 447 SVDVDSNGTDVLIGAPHYVEQTRGGQVSVCPPLRQGRARWQCDVAVLYGEQ-----498
Db 297 VVDI--NNDYTDLLVIGAPMYMDGPAIQ-----RWEAGAVVYVLYQNPDVGPGA 343
QY 499 -----GQPFGRFGAALTVLGDVNGDKLTDVAIGAPGEDNKGAVYLEHGTSGS 546
Db 344 SNRLSSTLIGGQIRSRFGLSIASIGDSNQDGFNDVAIGAPYEGDAGAVITYHG--SAN 402
QY 547 GISPSHSQRIAGSKLS-PRIQYFGQSISGGQDLTMDGLVDLTGQAQ--GHVLLRSQFVL 603
Db 403 GLKSTPAQVLTPTSLGHSGITTFGSLQGGQMDKNKVPDLLVGAESANTAVLIRTPVV 462
QY 604 RVKAIMFENPREVARNVFCNDQVVKKEAGEVRCVCHVQKSTRDRUREGQIOSV-----658
Db 463 SLDAITLNTPE-----IGINLENKTYE-LADGTMVTSFTAMT 497
QY 659 -----VTYDLALDSG-RPHSRAVENETKNSSTRQTVLGL-TQTCETILKQ 702
Db 498 CFYVTGNLYLPHDIDISYTVTVDSGLIANRRANFVDNDMSBITKTRRLAVSTQFCPLRAY 557
QY 703 LPNCIEDPVSPIVLRNFSVLGTPLSAFGN-----NDESRLQPHEILPIDMATMTQTKVSIQN 609
Db 558 VGNSEDKLTPIKVTLQYDL-----LRPVLAEADAQRLFTALPFPEK 753
QY 754 NCGNDNICODDLSITFSFMSLCLVGGPREFNVTVVRNDGEDSVRTQVTFPPDLDSY 813
Db 610 NCVN-NICYPDLDDVTV-PNLPIVIGTQQLTLDVSLNRRGEDAFQSSLSVYVPLGLQF 667
QY 814 RKVSTLQNRQSRWRLACESASSTEVSGAKSTCSINHPIFPEP-----SEVTENIT 867
Db 668 VRL-----ERKANMDFSVTCSESD-----LRIITCOTGNPMVGKNILFGLTLSTFQVS 717
```

```
QY 868 FDVDS-----KASLGNKLLKANVTSENMPRNTKTEFOLELPVKVAVYVMVTSHGVS TK 922
Db 718 GKXDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDOCTLLKLSASYPEIWMYSTQED 775
QY 923 YL-----NFTASENTSRVMOHOYOV-----SNLQORSILPIS-----LVFLVPV 960
Db 776 YVVPFPFPAKNASEADIGMEVM-HLVEVRNTGSSNAGEVSLNTQWPKQNEDEGEYLFYLLGI 834
QY 961 RLNQTVIMDRPQ-----VTFSENLSSTCHTKERLPSSHDFLAELRKAPVV 1005
Db 835 MTEEGVTCQLTQKANPEGVKLEPSTKAKLSNSTTQVSGRKRREPEVAALAQTDN--VI 892
QY 1006 NCSIAVCQRIQCDIPFGIOBEFNAT-----LKNLSFDWYIKTSHNHLIIIVSTAEILF 1059
Db 893 YCASDSCVLLINCTI-----DEINASKSKVVRILGRF--W-----ERTF 928
QY 1060 NDSVFTLLPGQGFVRSQSTETKVE--PFEVNP-----LP-----1092
Db 929 QKAVSELTPVVOATIASASAAVKTIPIYNIPLPRDFSOSTKASTLVTTELVPPVPTPIAW 988
QY 1093 -LIVGSSVGGLLLLALITAAALYKLGFFXR-----QYKDMMS--EGGPP 1132
Db 989 WIIIVSVLGGIILLIILGLMKCGFFERKKPGEEKYAPVASADKGGP 1038

Search completed: November 9, 2004, 12:19:28
Job time : 33.25 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5836.5	99.2	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4460	75.8	1153	1	ITAM_MOUSE	P05555 mus musculus
3	4371	74.3	1151	2	Q9J130	Q9i130 rattus norv
4	3799.5	64.6	920	2	Q28984	Q28984 sus scrofa
5	3469	59.0	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3401	57.8	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3319.5	56.4	1188	2	Q6KAS4	Q6kas4 mus musculus
8	3319.5	56.4	1188	2	BAD21383	Bad21383 mus muscu
9	3310.5	56.3	1169	1	ITAX_MOUSE	Q9qxb4 mus musculus
10	3236.5	55.0	1161	1	ITAD_RAT	Q9qyb7 rattus norv
11	1572	26.7	1165	1	ITAD_BOVIN	P61625 bos taurus
12	1547.5	26.3	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1546	26.3	1166	2	Q6TYB8	Q6tyb8 bos taurus
14	1546	26.3	1166	2	AAQ90015	Raq90015 bos tauru
15	1527.5	26.0	1163	1	ITAL_MOUSE	P24063 mus musculus
16	1523.5	25.9	1161	2	Q9WTV4	Q9wtv4 mus musculus
17	1513	25.7	1160	2	Q9R200	Q9r200 mus musculus
18	1395	23.7	1196	2	Q981F1	Q98tf1 cyprinus ca
19	1350.5	23.0	1086	2	Q96HB1	Q96hb1 homo sapien
20	1344.5	22.9	1187	2	Q98TF0	Q98tf0 cyprinus ca
21	1269	21.6	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1154.5	19.6	1167	2	O88340	O88340 rattus norv
23	1148.5	19.5	1167	1	ITAE_MOUSE	P60677 mus musculus
24	1140	19.4	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1109	18.8	1167	2	O88341	O88341 rattus norv
26	1093.5	18.6	1151	1	ITAI_HUMAN	P56199 homo sapien
27	1084.5	18.4	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1071	18.2	1170	1	ITAD_BOVIN	P53710 bos taurus
29	1057	18.0	1178	1	ITAE_MOUSE	Q62469 mus musculus
30	1057	18.0	1178	2	Q6PIC7	Q6pic7 mus musculus
31	1057	18.0	1178	2	AAH65139	Aah65139 mus muscu

32	1055	17.9	1180	1	ITAI_RAT	P18614 rattus norv
33	1054	17.9	1181	1	ITAI_HUMAN	P17301 homo sapien
34	1051.5	17.9	1167	1	ITAG_HUMAN	O75578 homo sapien
35	1043.5	17.7	1188	1	ITAH_MOUSE	P61622 mus musculus
36	1043.5	17.7	1188	2	O7TQC3	O7tcq3 mus musculus
37	1041	17.7	1171	2	O42094	O42094 gallus gall
38	1032	17.5	1038	2	Q8BS01	Q8bs01 mus musculus
39	1005	17.1	895	2	Q9WUF8	Q9wuf8 mus sp. itg
40	998.5	17.0	1160	2	Q8MKF4	Q8mkf4 felis silve
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	860	14.6	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	796	13.5	205	2	Q63001	Q63001 rattus norv
44	770.5	13.1	304	2	Q6PG66	Q6pg66 mus musculus
45	770.5	13.1	304	2	AAH57200	Aah57200 mus muscu

ALIGNMENTS

RESULT 1

ITAM_HUMAN						
ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.		
AC	P11215;					
DT	01-JUL-1989	(Rel. 11, Created)				
DT	01-OCT-1996	(Rel. 34, Last sequence update)				
DT	01-OCT-2004	(Rel. 45, Last annotation update)				
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)					
DE	(Neutrophil adherence receptor).					
GN	Name=ITGAM; Synonyms=CR3A, CD11B;					
OS	Homo sapiens (Human)					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88315033; PubMed=2457584;					
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;					
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";					
RT	J. Biol. Chem. 263:12403-12411(1988).					
RL	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88190151; PubMed=2833753;					
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;					
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol. chromosomal localization and homology to the alpha subunits of integrins.";					
RT	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).					
RL	[3]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88257215; PubMed=2454931;					
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;					
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";					
RL	J. Cell Biol. 106:2153-2158(1988).					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=93123748; PubMed=8419480;					
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;					
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";					
RT	J. Immunol. 150:480-490(1993).					
RL	[5]					
RN	SEQUENCE OF 9-1153 FROM N.A.					
RP	MEDLINE=8909893; PubMed=2563162;					
RX	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;					
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";					

RT "A partial genomic DNA clone for the alpha subunit of the mouse

FT	SIGNAL	1	16	
FT	CHAIN	17	1153	Integrin alpha-M.
FT	DOMAIN	17	1105	Extracellular (Potential).
FT	DOMAIN	1106	1129	Potential.
FT	DOMAIN	1130	1153	Cytoplasmic (Potential).
FT	REPEAT	31	84	FG-GAP 1.
FT	REPEAT	85	163	FG-GAP 2.
FT	DOMAIN	164	350	VMFA.
FT	REPEAT	337	400	FG-GAP 3.
FT	REPEAT	401	452	FG-GAP 4.
FT	REPEAT	454	515	FG-GAP 5.
FT	REPEAT	517	575	FG-GAP 6.
FT	REPEAT	580	632	FG-GAP 7.
FT	CA_BIND	465	473	Potential.
FT	CA_BIND	523	537	Potential.
FT	CA_BIND	592	600	Potential.
FT	SITE	1132	1136	GFRK motif.
FT	DISULFID	66	73	By similarity.
FT	DISULFID	105	123	By similarity.
FT	DISULFID	654	711	By similarity.
FT	DISULFID	770	776	By similarity.
FT	DISULFID	999	1023	By similarity.
FT	DISULFID	1028	1033	By similarity.
FT	CARBOHYD	58	58	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	86	86	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	391	391	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	734	734	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	772	772	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	801	801	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	881	881	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	907	907	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	941	941	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	980	980	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	994	994	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1022	1022	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1045	1045	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1051	1051	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1076	1076	N-linked (GlcNAc. .) (Potential).
FT	VARSPLIC	453	569	Missing (in isoform 2). /FTID=VSP_010473.
FT	CONFLICT	37	37	N->S (in Ref. 2).
FT	CONFLICT	683	683	V->G (in Ref. 2).
FT	SEQUENCE	1153	AA; 127480 MW; 178DB988AECB0343 CRC64;	
Query Match				
Best Local Similarity 75.8%; Score 4460; DB 1; Length 1153;				
Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;				
Qy	1	FNLDEHNTAFQENARGFGQSVVQLGSRVVVGAPQEIIVANORGSIVQCDYSTGSCBPI	60	
Db	17	FNLDEHPMTFQENAKFGQNVVQLGTSVVVAAPQAKAVNQTAIVQCDYSTSRCHPI	76	
Qy	61	RLQVPEAVNMSLGLSLAATSPQQLACGPTVHQTCSNTYVKGCLFLFGSNLRQDPOK	120	
Db	77	PLQVPPPEAVNMSLGLSLAVSTVPQQLACGPTVHQNCKENTYVNGLCVFLFGSNLRPPQ	136	
Qy	121	FPEALRGCPQEDSIAFLIDGSGIIPHDPRMKEWSTVMEQLKKSKTFLSLMQYSEEF	180	
Db	137	FPEALRGCPQEDSIVFLIDGSGIINNIDFQMKFEFVSTVMEQPKSKTLFLSLMQYSEF	196	
Qy	161	RIHFTKFFQNNPRSLIKPITOLLGRTHATGLRKVVRELFINITNGARKNAFKILFL	240	
Db	197	RIHFTFNDKKNPDRSHVSPIKQLNGRTKTASGIRKVVRELFKHTNGARENAKILWVI	256	
Qy	241	TDGEKFGDPLGYEDVIPELDREGVIRYVGVGDAPFRSEKSRQELNVTASKPRDHVFQIN	300	
Db	257	TDGEKFGDPLDYKDVIPEDRAGVIRYVGVGNAFNPKQSRRELDITASKPAGEHVFQD	316	
Qy	301	NFEALKTONLRKIKFIAEGTQTGSSSFHEHMSQEGFSAITSNGLLSTVGSYDWAG	360	
Qy	335	NTLVNLTQVIRFQTTQSTSEFHEHMSQEGFSAITSNGLLSTVGSYDWAG	376	

QY	361	G	V	F	L	T	S	K	E	K	S	T	I	N	N	T	R	V	D	S	D	M	N	A	I	G	Y	A	A	A	I	L	R	N	R	V	Q	S	L	V	I	G	A	P	Y	Q	H	I	G	V	A	M	F	R		420		
Db	377	G	A	F	L	T	S	K	D	K	V	F	I	N	T	R	V	D	S	D	M	N	A	I	G	Y	A	S	A	V	I	L	R	N	R	V	Q	S	L	V	I	G	A	P	Y	Q	H	I	G	V	A	M	F	R		436		
QY	421	Q	N	T	G	W	E	S	N	A	N	V	K	G	T	O	I	G	A	F	C	A	S	L	C	S	D	V	D	S	G	S	T	D	V	I	G	A	P	H	Y	Q	T	R	G	Q	V	S	V	C	P	L		480				
Db	437	E	N	F	G	T	W	E	P	H	T	S	I	K	G	S	O	I	G	S	Y	F	C	A	S	L	C	S	D	M	A	D	G	N	T	L	I	G	A	P	H	Y	Q	T	R	G	Q	V	S	V	C	P	L		496			
QY	481	P	R	G	O	R	A	M	O	C	D	A	V	L	Y	G	E	Q	O	P	R	G	A	L	T	V	L	D	V	N	G	O	K	I	T	D	V	A	I	G	A	P	E	E	D	N	R	C	A	V	L	F		540				
Db	497	P	R	G	-	R	A	R	M	O	C	E	A	L	L	H	G	D	Q	H	P	M	R	G	A	L	T	V	L	D	V	N	G	O	K	I	T	D	V	A	I	G	A	P	E	E	D	N	R	C	A	V	L	F		555		
QY	541	H	C	T	S	G	S	I	P	S	H	S	O	R	I	A	G	S	K	L	S	P	R	O	Y	F	Q	S	L	S	G	Q	D	I	T	M	O	G	L	V	D	L	T	V	G	A	O	H	V	L	L	R	S		600			
Db	556	Y	C	A	S	A	S	L	S	A	S	H	S	H	R	I	I	G	A	H	F	S	P	G	L	Y	F	Q	S	L	S	G	K	D	I	T	M	O	G	L	M	D	L	A	V	G	A	O	H	L	L	I	L	R	A	Q		615
QY	601	P	V	L	R	V	K	A	I	E	M	E	P	R	V	A	R	N	F	E	C	N	D	V	V	K	G	E	A	G	R	V	C	L	H	V	O	K	S	T	R	O	R	L	E	R	E	G	O	I	S	V	T		660			
Db	616	P	V	L	R	L	E	A	T	M	E	S	P	K	K	V	A	R	S	V	A	C	Q	O	V	L	K	N	D	A	G	E	R	V	C	L	R	V	R	N	K	T	O	R	L	E	G	I	O	S	V	T		675				
QY	661	Y	D	L	A	D	S	G	R	P	S	R	A	V	E	N	T	K	N	S	T	R	Q	T	O	V	L	G	I	T	O	T	C	E	L	K	Q	L	P	N	C	I	D	P	P	S	P	I	V	L	R	N	F		720			
Db	676	Y	D	L	A	D	P	V	R	S	I	R	A	P	F	E	T	K	N	T	R	R	T	O	V	F	G	L	M	O	K	E	T	L	K	I	L	P	D	C	V	D	S	V	P	I	L	R	N	Y		735						
QY	721	S	L	V	G	P	L	S	A	F	G	N	L	R	V	L	A	E	D	A	O	R	L	F	T	A	L	P	P	E	K	N	C	G	N	D	I	C	O	D	L	S	I	T	F	S	P	M	S	D	C	L	V	G		780		
Db	736	T	L	V	G	P	L	S	F	G	N	L	R	V	L	A	E	D	A	O	R	F	T	A	L	P	P	E	K	N	C	G	N	D	I	C	O	D	L	S	I	T	M	S	A	M	G	D	L	T	V	G		7				

```

RESULT 3
Q9J130
ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP Fathallah D.M. Sr., Zerria K. Jr.;
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AF268593; AAF81280.1; -
DR HSSP; P11215; 1BHQ.

```

DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 2.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS50234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match	74.3%;	Score 4371;	DB 2;	Length 1151;
Best Local Similarity	72.6%;	Pred. No. 3.5e-289;		
Matches	826;	Conservative 150;	Mismatches 159;	Indels 2; Gaps 2;
QY	1	FNLDTENAMTQENARGFGQSVOLQGSRRVVVGAPQETVAANORGSLYCCDYSTGSCPEI	60	
Db	17	FNLDTENPMTFQENASGFGQSVIQLGETRVVVAAPQEVKAVNQATGALYCCDYSTNRCDPI	76	
QY	61	RLOVPVEAVNMSLGLSLAAATSPPLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPOK	120	
Db	77	PLQVPPPEAVNMSLGLSLAAATVPPLLACGPTVHONCKENTYVNGLCYLFSGNLLRRPQQ	136	
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIHDFRPMKEWSTVMEQLKSKTFLPSLMQYSEBF	180	
Db	137	FPEALRGCPQESNIAFLIDGSGSINTIDFQKMKFEVSTVMDQFQSKTFLPSLMQYSEDF	196	
QY	181	RIHFTFEKFQNNPNRSLIKPIITQLLGRTHTATGLRKVVRLEFNIITGARKNAFKILFL	240	
Db	197	RTHFTFENDFKENPDKSHVRPIROLNGRTKTASGIRKVVRELFOKINGARDNAAKILVVI	256	
QY	241	TDGKFGDPLGYEDVIELDREGVIRVIVGVDAFRSEKSRQELINTVASKPPRDHVFQIN	300	
Db	257	TDGKFGDPLNYEDVIEPAEAEAGIIRVIVGVNAFHKPKQSRRELDITASKPAGDHVFQVD	316	
QY	301	NFEALKTIQNLREKIFAIETGTOGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	360	
Db	317	NFEALNIRNLQOEKIFAIETGTOGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	376	
QY	361	GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAAAIILERNVQSLVGLAPRYQHIGLVAMFR	420	
Db	377	GAFLYPSKDKASFINTTRIDSDMDNDAYLGYASAVISERNVQSLVGLAPRYQHIGLVAMFK	436	
QY	421	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVSCPL	480	
Db	437	QNFGEWEPHTDIIKGSQIGSYFGASLCSVDMDAGNTNLILIGAPHYETKRGQGVSCPL	496	
QY	481	PRGORARWQCDVLYGEOQPGWRFGAALTVLGDVNGDKLTVDVAIGAPGEDNRGAVILF	540	
Db	497	PRG-RARWQCEAILHGDQHPGWRFGAALTVLGDVNGDKLTVDVAIGAPGEQENQGVILF	555	
QY	541	HGTSGSGISPSHSQRIAGSKLSPLRYFGQSLSGGQDLTMDGHDVLTVAQGHVLLRSQ	600	
Db	556	HGASVASISTPHSORIAGARFSPGLYFGQSLSGGKDLTMDGLMDLVAQAQGRLLILRAQ	615	
QY	601	PVLVRKALIBFNPREVARNVFCNDOVVKGEAGEVVRVCLHVQKSTFRDLRERQIOSVVT	660	
Db	616	PVLRLKATWEFSFNKLARSVFACQEEVVRKNKQAGEVVRVCLQVRKNTKDLRERQIOSVVT	675	
QY	661	YDLALDSGRPHSAVFNETKNSRTRQTVLGLTQTCTETLKLQLPNCIEBPPVSPIVLRINF	720	
Db	676	YDLALDPGRSVVRAFFDETCKNGILRRIRVFLGTQKCTETKLILPDCVDNVSPIILRLNY	735	
QY	721	SLVGTPLSAFAGNLRPVLAEDAQBLFTALFPFPEKNCNGNDNICQDDLSITFSMSLDCLVVG	780	
Db	736	TLVGBEPLRSRDLRPVLAEMAQRIFTAMPFPFKNCGNDITCDDLSITVSSVTRNLTVVG	795	

```
QY 781 GPREFNVTVVRNDEGSDYRQVTFEPFLDLSYRKVSTLQNRQSGRWLACASASSTEV 840
Db 796 DSRDFDVSVTVLRNDEGSDYRQVTFEPFLDLSYRKVSTLQNRQSGRWLACASASSTEV 854
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 900
Db 855 QGVLSKSTIWDINHPIPFANSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 914
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 960
Db 915 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 974
QY 961 RLNQTVIWRDQVTFENSENLSSCTHKBRLPDSHDFLAELRKAPVNVCSIAVCQRIQCIP 1020
Db 975 QINKVTIWDPPQVTFENSENLSSCTHKBRLPDSHDFLAELRKAPVNVCSIAVCQRIQCIP 1034
QY 1021 PFGIOEENFATLKGNSLSDWIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSOTET 1080
Db 1035 SPNSKEIFNVTLQGNLLFDWYIETSHDHLILVSTAEILFNDSVFTLLPQCGAFVRSOTET 1094
QY 1081 KVEPEVENPVLPIVGVSSVGGILLALITAAIYKLGFFKQYKDMSEGGPPGAPQ 1137
Db 1095 KVEPYTVHNPVLIIVGVSSVGGILLALITAGLYKLGFFKQYKDMSEGGPPGAPQ 1151
```

RESULT 4

```
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE CD11b (fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS0234; VWEA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1
FT NON_TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.6%; Score 3799.5; DB 2; Length 920;
Best Local Similarity 78.8%; Pred. No. 3.2e-244;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;
QY 118 PQKFPALRGCPQEDSIAFLIDSGSIIIPDPRMKWVSTVMEQKKSKTLFSLMQYS 177
Db 1 PQKFPALRGCPQEDSIAFLIDSGSIIIPDPRMKWVSTVMEQKKSKTLFSLMQYS 60
QY 178 EEPRIHTFKEFQNNENPRSLIKPTQLGRTHATLGRVKEVLENIUNGAKNAFKIL 237
Db 61 EDFYTHTFDNRNPSKLLVRPIQLLGRTHATLGRVKEVLENIUNGAKNAFKIL 120
```

```
QY 238 FLITDGEKFGDPLGYEDVIVBELDREGVIRYVIGVDAPRSEKSEQELNTVASKPRDHVF 297
Db 121 VVITDGEKFGDPLGYEDVIVBELDREGVIRYVIGVDAPRSEKSEQELNTVASKPRDHVF 180
QY 298 QINNFEALKITONOLREKIIFAEICTOTGSSSSFFHEMSQEGFSAITNSGPLLSTVGSYD 357
Db 181 QVNNFEAVTKITONOLREKIIFAEICTOTGSSSSFFHEMSQEGFSAITNSGPLLSTVGSYD 240
QY 358 WAGGVLYTSKEKSTFNTMTVDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVA 417
Db 241 WAGGAFLEHMPKDRVIFNTTEVSDMDNDAYLGAAVEILNRQAQSLVLGAPRYQHTGLW 300
QY 418 MFRQNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDLVLGAPHYIYQTRGGQVSV 477
Db 301 MFKQNSGAWKNAIDKISQIGSYFGASLCSVDVDRDSSDLVLGAPHYIYQTRGGQVSV 360
QY 478 CPLPRGQARQWQCDVLYGEOGPWGRGAALTVLGVNGDKLTDVAIGAPGEEDNRGAV 537
Db 361 CPLPQG-RAKWQCRVILGCEGQHPWSRGAALTALGVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLEHGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLL 597
Db 420 YLEHGTSELGLSPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVLL 479
QY 598 RSQPVLRVKAIMEFNPREVARNVPECNDQVVGKEAGVRVCLLVKQKSTRDLRREGQIQS 657
Db 480 RSQPVLRVKAIMEFNPREVARNVPECNDQVVGKEAGVRVCLLVKQKSTRDLRREGQIQS 539
QY 658 VVTVDLALDSGRPHSRVFNETHKSTRQTVLGJTOTCETLKLQLPNCIEDPVPVILR 717
Db 540 IITVDLALDPGRPHSRVFNETHKSTRQTVLGJTOTCETLKLQLPNCIEDPVPVILR 599
QY 718 LNFSLVGTPLGAFGNLRPVLAEADQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCL 777
Db 600 LNFSLVGTPLGAFGNLRPVLAEADQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCL 659
QY 778 VVGGRPRENVTVVRNDEGSDYRQVTFEPFLDLSYRKVSTLQNRQSGRWLACASASSTEV 837
Db 660 VVGGRPRENVTVVRNDEGSDYRQVTFEPFLDLSYRKVSTLQNRQSGRWLACASASSTEV 719
QY 838 TEVSGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 897
Db 720 TEESTALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 779
QY 898 TEFQLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 957
Db 780 TEFQLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 839
QY 958 VVRLNQTVIWRDQVTFENSENLSSCTHKBRLPDSHDFLAELRKAPVNVCSIAVCQRIQC 1017
Db 840 VVRLNQTVIWRDQVTFENSENLSSCTHKBRLPDSHDFLAELRKAPVNVCSIAVCQRIQC 899
QY 1018 DIPFGIOEENFATLKGNSLSDWIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSOTET 1080
Db 900 DIPFGIOEENFATLKGNSLSDWIKTSHNHLILVSTAEILFNDSVFTLLPQCGAFVRSOTET 840
```

RESULT 5

```
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702; Q81V66;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Lew M5).
GN Name=ITGAx; Synonyms=CD11c;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```


RX MEDLINE=88166645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "cDNA cloning and complete primary structure of the alpha subunit of a
 RT leukocyte adhesion glycoprotein, p150,95";
 RL EMO J. 6:4023-4028(1987).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
 RT molecule";
 RL J. Biol. Chem. 265:2782-2788(1990).
 RN [3]
 RN ERRATUM.
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751(1990).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Ustin T.B., Toshitsuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Casavant T.L., Scheetz T.E.,
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RN SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RT and p150,95 leukocyte adhesion proteins";
 RL J. Immunol. 138:2381-2383(1987).
 CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis.
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 CC granulocytes.
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81695; AAA51620.1; --
 DR EMBL; M29165; --; NOT ANNOTATED CDS.
 DR EMBL; M29487; AAA51620.1; ALT SEQ.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29483; AAA51620.1; JOINED.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29485; AAA51620.1; JOINED.
 DR EMBL; M29486; AAA51620.1; JOINED.
 DR EMBL; BC038237; AAH38237.1; --.
 DR PIR; A36584; RWHUIC.
 DR PDB; 1N3Y; X-ray; A=141-338.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; --.
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0009887; P:organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWFA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SMO0191; Int alpha; 5.
 DR SMART; SMO0327; VWFA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 19
 FT CHAIN 20 1163 Integrin alpha-X.
 FT DOMAIN 20 1107 Extracellular (Potential).
 FT TRANSMEM 1108 1128 Potential.
 FT DOMAIN 1129 1163 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 ? FG-GAP 2.
 FT DOMAIN 165 351 VWFA.
 FT REPEAT ? 401 FG-GAP 3.
 FT REPEAT 402 453 FG-GAP 4.
 FT REPEAT 455 517 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 466 474 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1131 1135 GPFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 655 712 By similarity.
 FT DISULFID 771 777 By similarity.
 FT DISULFID 848 863 By similarity.
 FT DISULFID 998 1022 By similarity.
 FT DISULFID 1027 1032 By similarity.
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).
 FT VARIANT 48 48 W -> R (in dbSNP:11574633).
 FT FTID=VAR 018672.
 FT CONFLICT 209 209 T -> S (in ref. 4).
 FT CONFLICT 251 251 T -> A (in ref. 4).
 FT CONFLICT 469 469 T -> S (in ref. 4).
 FT CONFLICT 490 490 G -> A (in ref. 2).
 FT CONFLICT 547 547 E -> K (in ref. 4).
 FT CONFLICT 756 756 D -> L (in ref. 1).
 FT CONFLICT 819 819 I -> V (in ref. 4).
 FT CONFLICT 1161 1163 SEK -> TPHYQDNV (in ref. 4).
 FT STRAND 150 157

RL Gene 171:291-294 (1996).
 RP [4]
 RX INTERACTION WITH VCAM1.
 RA MEDLINE=99059842; PubMed=9841932;
 RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;
 RT "alphabeta2 integrin is expressed on human eosinophils and functions
 RT as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
 RT 1).";
 RL J. Exp. Med. 188:2187-2191 (1998).
 RP [5]
 RX INTERACTION WITH VCAM1.
 RA MEDLINE=99370002; PubMed=10438935;
 RA Van der Vieren M., Crowe D.T., Hoekstra D., Vazquez R., Hoffman P.A.,
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
 RT binding interface between I domain and VCAM-1.";
 RL J. Immunol. 163:1984-1990 (1999).
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
 CC VCAM1. May play a role in the atherosclerotic process such as
 CC clearing lipoproteins from plaques and in phagocytosis of blood-
 CC borne pathogens, particulate matter, and senescent erythrocytes
 CC from the blood.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
 CC associates with beta-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
 CC lines and subsets of peripheral blood leukocytes and strongly on
 CC tissue-specialized cells, including macrophages foam cells within
 CC atherosclerotic plaques, and on splenic red pulp macrophages.
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL; U37028; AAB38547.1; -
 DR EMBL; U40274; AAB60634.1; -
 DR EMBL; U40275; AAB60635.1; -
 DR EMBL; U40276; AAB60636.1; -
 DR EMBL; U40277; AAB60637.1; -
 DR EMBL; U40278; AAB60638.1; -
 DR EMBL; AF187881; AAF62875.1; -
 DR HSSP; P11215; 1BHQ.
 DR Genew; HGNC:6146; ITGAD.
 DR MIM; 602453; -
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0016337; P:cell-cell adhesion; NAS.
 DR GO; GO:0007160; P:cell-matrix adhesion; NAS.
 DR GO; GO:0006955; P:immune response; NAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWFA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; WVA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 DR Calcium; Cell adhesion; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 17 Potential.

FT CHAIN 18 1162 Integrin alpha-D.
 FT DOMAIN 18 1100 Extracellular (Potential).
 FT TRANSMEM 1101 1121 Potential.
 FT DOMAIN 1122 1162 Cytoplasmic (Potential).
 FT REPEAT 32 85 FG-GAP 1.
 FT REPEAT 86 ? FG-GAP 2.
 FT DOMAIN 150 332 VWFA.
 FT REPEAT 350 400 FG-GAP 3.
 FT REPEAT 401 452 FG-GAP 4.
 FT REPEAT 454 516 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 465 473 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1127 1131 GPEKR motif.
 FT DISULFID 67 74 By similarity.
 FT DISULFID 106 124 By similarity.
 FT DISULFID 655 710 By similarity.
 FT DISULFID 769 775 By similarity.
 FT DISULFID 846 861 By similarity.
 FT DISULFID 994 1018 By similarity.
 FT DISULFID 1023 1028 By similarity.
 FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 500 500 Missing (in Ref. 2).
 FT CONFLICT 515 518 GHPW -> ATP (in Ref. 2).
 FT CONFLICT 825 825 L -> V (in Ref. 2).
 FT CONFLICT 984 984 V -> A (in Ref. 2).
 SQ SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;

Query Match 57.8%; Score 3401; DB 1; Length 1162;
 Best Local Similarity 59.3%; Pred. No. 1.8e-217;
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;
 QY 1 FNLDTENAMTFQENARFGQSVVLQSGRSVVGVAPOEIVAAQNGSLYQCDYSTGSGCEPI 60
 DB 18 FNLDVBEPTIFQEDAGGFGQSVVQFGGSLRVVGLVVAQNGTGLYDCAATGMCQPI 77
 QY 61 RLQVPVEAVNMSLGLAATTSPPOLLACGPTVHTCSENYVKGICFLFGSNLRQOPQK 120
 DB 78 PLHIREAVNMSLGLTAASTNGSRLACGPTLHRCVGENSYSGKSLGSRW-ETIQT 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKEWSTVMEOLKSKTLFSLMOYSEEF 180
 DB 137 VPDATPECPHQEMDIVFLIDGSGSIDQDNFNQMGFQVAVMGQFGTDTLFAIMOYSNL 196
 QY 181 RHFTFKFQNNPNPSLKKPTQLLGRTHATGLKVVVRELNFITNGARKNAKILFL 240
 DB 197 KHTFTPTQRTSPSQSLVDPIVLQKGLTFTATGILTVTQLFHHKNGARKSAKLLI 256
 QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVIGVDGAFRSEKSRQELNTVAKSPRDRHV 300
 DB 257 TDGQKYKPLEYSDVIPAQKAGIIRYAIQGVGHAFQPTARQELNTISSAPQDHVKVD 316
 QY 301 NFEALKTIQNLREKIFATGQTGSSSFEHMSQEGFSAITSNGPLSTVSGYDNAG 360
 DB 317 NFAALGSIQKLOEKIYAVEGTQSASSSFQHEMSQEGFSTALTDGLFLGAVFSWSG 376
 QY 361 GVFLYTSKEKSTFTINMTRVDSMDNDAYLGAAAIILRNVRQSLVGLGAPYQHGLVAMFR 420
 DB 377 GAFLYPPNMSPTFINNSQENVDMRDSYLGYSTELALWKGQNLVGLGAPYQHTKAVIT 436
 QY 421 QNTGWNESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYEQTRGGQSV 480
 DB 437 QVSRQWRKKAETVGTQIGSYFGASLCSDVDSDGSTDLLIGAPHYEQTRGGQSV 496

481 PRQARWQCDVAVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
497 PRQORVQWQCDVAVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 600
557 HGASEGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 616
601 PVLRVKALMEFNPREVARNVPECDNVQVVKGEAGEVVRVCHVQKSTRDRLEGGQISVVT 660
617 PVLKVGVMARFSPVEVAKAVYRCWEKPSALEAGDAVCLTIQKSSLDQL-GDIQSVSR 674
661 YDLALDGRPHSRVAFNETKNSRTRQVGLGTOTCTETLKQLPNCIEDPVSPIVRLNF 720
675 FDLALDGRHLSTRAIFKTKTRALTRVKTGLNKHCHESVKLLLPACVEDSVPTLLRLNF 734
721 SLVGTPLSAFGLNLPVLAEDAQRFTALFPFEKNCNDNICODDLSITFFSMSLDCLVVG 780
735 SLVREPISPQNLRLPVLAVGSDQLFTASLPFEKNCQDGLCEGLGVTLSPGQLTLVG 794
781 GPREFNVTVVRNDEGDSYRTQVTPFFPLDLSYRKVSTLQNRQSRWELACESASSTEV 840
795 SSLELNVITVWVNGEDSYGTVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853
841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRINKTEF 900
854 EG-LRSSRCVNHPIFHEGNSGTFTVTFDVSFKATLGRMLMRASSENKASSKATF 912
901 QLELPVKAVYVMTSHGVSUKYLNFTASENTSRVMOHVOYVNLGORSIPISVFLVP 959
913 QLELPVKAVYVMTSHGVSUKYLNFTASENTSRVMOHVOYVNLGORSIPISVFLVP 972
960 VELNQTIVTWDPRQVTFSENLSTCHTKERLPKSHSDFLAELRKPAPVNCISAVCQRIQDI 1019
973 VLLNGVAVWVWVMEAPSQSL-PCVSEKPPQHSDFLTQISRSPLDCSADIACLPQCDV 1030
1020 PFFGIQEEFNATLKNLSPDWIKYKSHNHLIVSTABILFNDSVFTLLPGQAFVRSQFE 1079
1031 PSFSVQEEDELFTLKNLSPGWRETLQKVLVSVVAEITFDTSVVSQLPQGEAFMRAQME 1090
1080 TKVEFEVNPPLPLVINGSSVGLLALALATLALYKLGFFKROVKDMMSE 1128
1091 MVLEEDVYNALPIINGSSVGLLALALATLALYKLGFFKROVKDMMSE 1139

RESULT 7
Q6KAS4
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN Name=MFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RA "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries";
RL DNA Res. 11:167-180(2004).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AK131133; BAD21383.1; -
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.

InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS00234; VWEA; 1.
DR Cell adhesion; Integrin; Transmembrane.
KW NON TER 1
FT SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;
Query Match 56.4%; Score 3319.5; DB 2; Length 1188;
Best Local Similarity 56.8%; Pred. No. 4.9e-212;
Matches 647; Conservative 171; Mismatches 303; Indels 19; Gaps 7;
QY 1 FNLDTENAMTPOENARGFGQSVVQLQGSRVVVGAPQETVAANORGSLYQCDYSTGSCPEI 60
DB 39 FNLDAEKPFTFHMDDGAEFGHSLVQYDSSVWVVGAPKEIKATNQIGGLYKCGYHTGNCPEI 98
QY 61 RLOYPVAVNMISLGLSLAATTPPQLLACGPTVHTQTSENTYVKGLCFGLFNSLRQQPK 120
DB 99 SLOVPPPEAVNMISLGLSLAATNPWLLACGPTVHTQTSENTYVKGLCFGLFNSLRQQPK 157
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDPRMKEWSTVMEQLKSKTSLFSLMOYSEF 180
DB 158 FPTAQCEPKQDDIVFLIDGSGSISSTDFEKLMDFKAVMSQLQRPSTFSLMQFSDF 217
QY 181 RIHFTFKFQNNPNRSLIKITQLLGRTHATGLRKVVRELENTGARKNAKILFL 240
DB 218 RVHFTFNNFISSTPSLSLSDSVRQLRGYTYTASAIKHVITELFTTQSGARQATKVLVI 277
QY 241 TDGEKFGPLGVEDYVIBELDREGVIRYVIGVDAPFRSKSRQELNTVASKPRDHFQIN 300
DB 278 TDGRKQGNLSYDSVIMPAEASIIRYAIGVOKAFYNEHSKQELKAIASMPSEHYVFSVE 337
QY 301 NPEALKTTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNGPILLSVSGVDWAG 360
DB 338 NFDALKDIEQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTDPGVLGAVGSFWSG 397
QY 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGAAAILRNVRQSVILGAPVQHGLVAMFR 420
DB 398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAPWKGVHSLILGAPRHQHTGVIFT 457
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGGQVSCPL 480
DB 458 QESRHRWRPKSEVRGTQIGSYFGASLCSVDMDRDSGLDVLIGVPHYEHTRGQVSCPM 517
QY 481 PRQARWQCDVAVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 518 P-GVGRWHCGTTLHGEGHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVYIF 576
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 600
DB 577 HGASEGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHVLLRSQ 636
QY 601 PVLRVKALMEFNPREVARNVPECDNVQVVKGEAGEVVRVCHVQKSTRDRLEGGQISVVT 660
DB 637 PIURVPTVHTFTFAEISRSVFECQVAPETLSDATVCLHIESPKTL-GDLRSTVT 694
QY 661 YDLALDGRPHSRVAFNETKNSRTRQVGLGTOTCTETLKQLPNCIEDPVSPIVRLNF 720
DB 695 FDLALDGRHLSTRAIFKTKTRALTRVKTGLNKHCHESVKLLLPACVEDSVPTLLRLNF 754
QY 721 SLVGTPLSAFGLNLPVLAEDAQRFTALFPFEKNCNDNICODDLSITFFSMSLDCLVVG 780
DB 755 SLVGVPISSQLNQLPMLAVDDQTYFTASLPFEKNCNADHICQDDLSVVFGLDKTLVVG 814
QY 781 GPREFNVTVVRNDEGDSYRTQVTPFFPLDLSYRKVSTLQ-NQSRQSWR 829
DB 815 SDLELNVITVWVNGEDSYGTITVTFVVGSLPFRVAGQVFLRKCKEDQQRGQHSIH 874

Query Match	56.4%;	Score	3319.5;	DB 2;	Length	1188;			
Best Local Similarity	56.8%;	Pred.	No. 4.9e-212;						
Matches	647;	Conservative	171;	Mismatches	303;	Indels	19;	Gaps	7;
QY	1	FNLDTENAMTFQENARGFCQSVVOIQGSRVVVGAPQETVAANQROGSLYQCCDYSTGSCPEI	60						
Db	39	FNLDAEKPTHFHMUGAEGFHSVLQYDSSWVVVGAPKEIKATNIQIGLYKCGYHTGNCPEI	98						
QY	61	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHOTCSENTVYVVGKLCFLPSGNLRQPOK	120						
Db	99	SLQVPPEAVNMSLGLSLAATNPMSLLACGPTVHHTCRENIYTLGCLPILLSSSFKQS-QN	157						
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMNEQLKKSKTFLSLMQYSEEF	180						
Db	158	FTAQOCEKQPODQVFLIDGSGSISSTDFEKMILDVKAVMSQLQRPSTFSLMQPSDYF	217						
QY	181	RLHFTFKCFQNNPNRSILKIPITQLLGRTHATGLRKVVRELFNITNGARKNAKPIFLFL	240						
Db	218	RVHFTFNFISSPLSLDVSRLQURGYTATSAIKKHVITELFTQTSGARQDATKVLIVI	277						
QY	241	TDGEYFGDPLGVEDVIPELDREGVIRYVIGVDFAFRSEKSRQELNVTWASPPRDHVFQIN	300						
Db	278	TDGRKQGNLSYDSVIPMAEASIIRYAIGVKAFYNEHSKQELKATASMPHSHEYFVSVE	337						
QY	301	NPEALKTTIONLREKIPALEGTQTCGSSSSFHEMSQGFSAAITNSGPLLSTVGSYDWAQ	360						

RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,
 RA Tsuchiya H.:
 RT "Isolation of genes selectively expressed by dendritic cells."
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibronogen. It
 CC recognizes the sequence G-P-R in fibronogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis (By similarity).
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; AF211864; AAF23492.1; --
 DR HSSP; P20702; I1N3Y.
 DR MGD; MGI:96609; Ilgax.
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWF; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWF; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 DR Calcium; Cell adhesion; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 KW SIGNAL 1 19 Potential.
 FT CHAIN 20 1169 Integrin alpha-X.
 FT DOMAIN 20 1116 Extracellular (Potential).
 FT TRANSMEM 1117 1137 Potential.
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 ? FG-GAP 2.
 FT REPEAT 152 330 VWFA.
 FT DOMAIN 152 330 FG-GAP 3.
 FT REPEAT ? 402 FG-GAP 4.
 FT REPEAT 403 454 FG-GAP 5.
 FT REPEAT 456 518 FG-GAP 6.
 FT REPEAT 519 577 FG-GAP 7.
 FT REPEAT 582 634 Potential.
 FT CA_BIND 467 475 Potential.
 FT CA_BIND 531 539 Potential.
 FT CA_BIND 594 602 Potential.
 FT SITE 1140 1144 GEFKR motif.
 FT SITE 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 656 711 By similarity.
 FT DISULFID 770 776 By similarity.
 FT DISULFID 858 873 By similarity.
 FT DISULFID 1007 1031 By similarity.
 FT DISULFID 1036 1041 By similarity.
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 267 267 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 949 949 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1059 1059 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1084 1084 N-linked (GlcNAc...) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033219A6 CRC64;
 Query Match 56.3%; Score 3310.5; DB 1; Length 1169;
 Best Local Similarity 56.6%; Pred. No. 1.9e-211;
 Matches 645; Conservative 173; Mismatches 303; Indels 19; Gaps 7;
 QY 1 FNLDTENAMTFQENARGFGSGVVLQGSRRVVVVGAPQELVAANORGSYLQCDYSTGSCBPI 60
 DB 20 FNLDIAEKLTTHPMGDAEFGHSLVLYQSDSSVVVVGAPKEIKATNQIGGLYKCYGHTGCEPI 79
 QY 61 RLQVPVAVNMGLSLAATTPPQLLACGPTVHTCTSENTYVYKGLCFLFGSNLRQQPQK 120
 DB 80 SLQVPEPAVNLISGLSLAATNPWLLACGPTVHTCTSENTYVYKGLCFLFGSNLRQQPQK 138
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDFRMEKWEVSTVMEQLKSKTSLFSLMOYSEEF 180
 DB 139 PPTAQCEPCQDQDIVFLIDGSGSISSTDFEKLDFKAVMSQLQRPSTFSLMOYSEEF 198
 QY 181 RIHFTFKFQNNPNSRLIKITQLLGHHTATGLKVVRELFNITGARNNAKILFLL 240
 DB 199 RVHFTFNNFISTSSPLSLGSRVQLRGVYTTASAIKHVITELFTTQSGARQDATKVLVI 258
 QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVGDAFRSEKSKQELNVTASKPRDHVQIN 300
 DB 259 TDGKQGNLISYDSVIPMAEASIIRYAIGVGKAFYNEHSKQELKAIASMPGSHYVFSVE 318
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPEHEMSQEGFSAITNGPLLSVTGSDVWAG 360
 DB 319 NFDALKDIENQLEKIFAIEGTETPSSSTFELEMSQEGFSAVTFDGPVLGAVGSPWSG 378
 QY 361 GYFLYTSKEKSTFINMTRVDSMDNDAIYGAAAILLRNVQSLVLGAPYQHIGLVAMFR 420
 DB 379 GAFLYPSNMRTPTFNNMQENEDMRDAYLGYSTALAPFWKGVHSLILGAPRHQTKGVIFT 438
 QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLIGAPHYYEYTRGQGVSCPL 480
 DB 439 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDVLVLIGVPHYYEYTRGQGVSCPM 498
 QY 481 PRGQARWQCDVLYGEQQQPMGRFGAALTGLVDVNGDKLTVAICAPGEEDNRGAVYLF 540
 DB 499 P-GVGRWHCGTTLHGEQHPWGRFGAALTGLVDVNGDSLADVAICAPGEEDNRGAVYIF 557
 QY 541 HGTSGSGISPSHQRIAGSLPRLQYFGQSLGSGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 DB 558 HGASRQDIAPSPQSRIASQIPSRIQYFGQSLGSGQDLTRDGLVDLAVGSKRVLLRTR 617
 QY 601 PVLVRKAIMEFNPREVARNVFCNDQVWKGKAGEVRVCLHVOKSTRDLRLEGIOQSVIT 660
 DB 618 PILRVSTVHTFTAEISRSVFECQEQVAPQTLSDATVCLHIESPKTL--GDLRSTVT 675
 QY 661 YDLALDSGRPHRAVFNETKNSSTRQTVLGLTQCTETKLQLPNCIEDPSPVLRINF 720
 DB 676 FDLALDHGRLSTRAIFKETKTRALTRVKTGLNKHCSVKLLLPACVEDSVPTITLRNF 735
 QY 721 SLVGTPLSAGNLPPVLAEDAQRLLFTALFPKKNCGNDNICODDLSITFSFMSLDCLVVG 780
 DB 736 SLVGVPISSLQNLQPLAVDDQTYTASLPKKNCGADHICQDLSVVVFGPDKTLVVG 795
 QY 781 GPREFNVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQ-----NQSRQSWR 829
 DB 796 SDLEINVDVTVSNDEGDSYGTVTILFYVGLSFRFVABGQVFLRKKEDQQRGQHSILH 855
 QY 830 LACESASSTEVSGALKSTSCSINHPIPEENSEVFNITFDVDSKASLGNKLLKANTVSE 889
 DB 856 LMCD--STPDRSQGLWSTSCSRHVFRRGSGMTFLVTFDVSXPAELGDRLLRLARVGE 913
 QY 890 NNMPTNTEFOLELPVKAIVYVMVTSEGVSTKYLNFTASE--NTSRVMQHOYQVSNLGR 948
 DB 914 NNVPCTPKTTFQLELPVKAIVYVMVTSEGVSTKYLNFTASE--NTSRVMQHOYQVSNLGR 973
 QY 949 SIPLISFLVPLVRLNQTIVWRPQVTFSESNLSSTCHTKERLPSHSDFLAELRKA PVNCS 1008
 DB 974 DVEFSINFWPIELKGEAVW--TVMSVHPQNPLTQCYVRNRLKPTQFDLLTHMQKSPVLDGS 1032

QY 1009 IAVQRTQCDIPFGIOEFENATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLP 1068
Dd 1033 IADCLHRCIDPSIGLDLYFLKGNLSFGWISQTLQKKVLLSEAEITFNTSVISQLP 1092
QY 1069 GQGAFTVSQETKVEPEVENPLIVGSSVGGLLALITAAALYKLGFKRQYKDMWSE 1128
Dd 1093 GQEAFLRAQTKTLEMYKHNPLIVGSSVGGLLALITAILYKAGFKRQYKEMLEE 1152

RESULT 10
ITAD RAT
ID ITAD RAT STANDARD; PRT; 1161 AA.
AC OSQYER;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley;
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin."
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF021334; AAF21241.1; --
DR HSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWFA.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWFA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.

FT REPEAT 88 ?
FT DOMAIN 152 334 FG-GAP 2.
FT REPEAT 352 402 VWFA.
FT REPEAT 403 454 FG-GAP 3.
FT REPEAT 456 517 FG-GAP 4.
FT REPEAT 519 577 FG-GAP 5.
FT REPEAT 582 634 FG-GAP 6.
FT CA_BIND 467 475 FG-GAP 7.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GFPR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.0%; Score 3236.5; DB 1; Length 1161;
Best Local Similarity 57.5%; Pred. No. 1.6e-206;
Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLDTENAMTFQENARFGOSVVOVGSRVVGAPQEIIVAAQNGSLYOCYVSTGSCPIR 61
Dd 21 NLDVEEPTVFREDAAFGQTVVQFGSRLVVGAPLEAVAVNQTGLYDCAPATGMCPIV 80
QY 62 LQVPVEAVNMSLGLSLAATTPPOLLACGPTVHPTCSENTYVKGICFLFGSNLRQPOKPF 121
Dd 81 LRSPLAEVNMISLGLSLVATNNAQLACGPTAQRACVKNYAKGSCLLGSSL-QFIQAV 139
QY 122 PEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMWSTVMEQLKKSKTLFISIMOVSEBR 181
Dd 140 PASMPECPQEMDIAFLIDGSGSINQDFAQMKDFVXALMGFEASTSLFSLAQYSNILK 199
QY 182 IHFTFKFQNNPNRSLIKPTTOLLGTHATGKRVVRELFTNIGARKNAKILFLLT 241
Dd 200 THFTTFEKNLLDQSLVDPIVQOGUTYATGRTVMEELFHSKNGSRKSAKKILLVIT 259
QY 242 DGEKFGDPLGYEDVPIPELDREGVIRYVIGVDAPFRSEKSRQELNTVASKPDRHVFQINN 301
Dd 260 DGQYRDPLEYSDVI PAADKAGIIRYALVGDAFQEP TALKELNTIGSAPDQDHVKVGN 319
QY 302 FEALKTIONLRKIFAIEGTQTGSSSFHEMSQEGFSAATSNGLPLSTVSGSYDWAG 361
Dd 320 FAALRSIQRLQEKIFAIEGTQSRSSSFQHEMSQEGFSSALTSDFVLGVNGSFSWGG 379
QY 362 VFLYTSKEKSTFTNMTVDSMDNDAYLGAAAILNRNVOSLVLGAPRYOHIGLVAMFRQ 421
Dd 380 AFLYPNTRPTFTNMSQENVMDRSDVILGYSTAVAFWKGVHSLILGAPRHQTKGVITQ 439
QY 422 NTQMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLIGAPHYVTEOTGGQSVCLPL 481
Dd 440 EARHWRPKSEVRGTQIGSVFGASLCSVDVDRDGSSTDLVLIGAPHYVTEOTGGQSVFPVP 499
QY 482 RQORARWQCDAVLYGQGFQWGFGAALTIVLGDVNGDKLTDVAIGAPGEDNNGAVYLFH 541
Dd 500 -GVRGEWQCEATLHGQCHQWGRFVALTVLGDVNGDNLDVAIGAPGEESRGAVYIFH 558
QY 542 GTSGSGISPSHQRINGSKLSRPLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLRLSQP 601
Dd 559 GASRLIMPSPQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAGQHVLRLSLP 618

QY	602	VLRYKAIMENPREVARNVFCNDQVKGKAGEVRVCLHVQKSTRDLRGOIQSVVY	661
Db	619	LLKVELSIRFAPMEVAKAVYQCWERTPTVLEAGETVCLTVHKGSPDLL--GNVQGSVRY	676
QY	662	DLALDSGPHSRAVNETKSTRRTQVGLTQTCETLKLOLPCNIEDPVPSPVLRNLFS	721
Db	677	DLALDPGLRLISRAIDETKNCITLGRKTLGLGDHCECTVKLLPDCDVEDAVSPVLRNLFS	736
QY	722	LVTPLSAFNGLRPLVLAEDAQRLFTALPFPKNCNDNICQDDLSITFSFMSLDCLVVGG	781
Db	737	LVDRASAP--RNLPVLAVGSQDHITASLPFKNCKQELLCEGLGIFSNFSGLOVLVVG	795
QY	782	PREFNVTVRNDEGDSYRTQVTFPPDLDSYKYSTLTQNRQSRWRLACASASTVS	841
Db	796	SPELVITVWNEGEDSYGLVKFYFAGLSYRRVTGTQ-QPHQYPLRLACAEFAAQED	854
QY	842	GALKSTSCSINHPIIPENSEVEFNTFDVDSKASLGNKLLKANVTSENNMERTNKTETQ	901
Db	855	--LRSSSCSINHPIIFREGAKTTFMIFDVSYKAFGLGDRLLLRKAKASSENKPDNTKTAQ	912
QY	902	LPLPVKAYVMVTSHGVTCKYLNFTASENTSR--VMQHOYQVSNLQORSLPISLVLVPV	960
Db	913	LPLPVKITYTLLISROEDSTNVHVFSSHGRRQEAHRYVNNLSPLKLAVRVNEWV	972
QY	961	RLNQTVIMDRPQVTFSENLS--TCHTKERLPSSHDFLAELKAPVNCVSIANCQICQDI	1019
Db	973	LLNGVAVMD--VTLSSPAQGVSCVSMKPPQNPDELTOIQRSVLDCSIADCLHPRCDI	1029
QY	1020	PFGIQEENATLKNLSFDWYIKTSNHLIIYSTAELFNDVSFTLLPGQAFVRSQTE	1079
Db	1030	PSLIDIODELDFILRNLFSQVSGTQKQKLVVSEAEITFDTSVYQLPQOEAFIAQVE	1089
QY	1080	TKYPEPEVNPPLIVGSSVGGILLALLIITAALYKLGFFKRYQKMDM	1126
Db	1090	TTLLEYVYVEPIFLVAGSSVGGILLALLIITVLYKLGFFKRYQKEML	1136
RESULT 11			
ID	ITAL BOVIN	STANDARD;	PRT; 1165 AA.
AC	P61625; 2004 (Rel. 44, Created)		
DT	05-JUL-2004 (Rel. 44, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1		
DE	alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha		
DE	chain) (CD11a).		
GN	Name=ITGAL;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	PubMed=14697514; DOI=10.1016/j.j.gene.2003.09.043;		
RA	Fett T., Zecchinon L., Baise E., Desmecht D.;		
RT	"The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,		
RT	characterisation and comparison with the human and murine		
RT	glycoproteins."		
RL	Gene 325:97-101(2004).		
CC	-!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,		
CC	ICAM3 and ICAM4. It is involved in a variety of immune phenomena		
CC	including leukocyte-endothelial cell interaction, cytotoxic T-cell		
CC	mediated killing, and antibody dependent killing by granulocytes		
CC	and monocytes (By similarity).		
CC	-!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L		
CC	associates with beta-2 (By similarity).		
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
CC	-!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins		
CC	with I-domains do not undergo protease cleavage.		
CC	-!- SIMILARITY: Belongs to the integrin alpha chain family.		

CC	-!- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-!- SIMILARITY: Contains 1 VWFA domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	EMBL; AY267467; AAP94035.1; --		
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.		
DR	PROSITE; PS50234; VWFA; 1.		
KW	Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;		
KW	Repeat; Signal; Transmembrane.		
FT	SIGNAL	1	23
FT	CHAIN	24	1165
FT	DOMAIN	24	1084
FT	TRANSMEM	1085	1105
FT	DOMAIN	1106	1165
FT	REPEAT	40	88
FT	REPEAT	89	146
FT	REPEAT	153	324
FT	DOMAIN	347	397
FT	REPEAT	398	452
FT	REPEAT	454	513
FT	REPEAT	515	572
FT	REPEAT	575	627
FT	CA_BIND	465	473
FT	CA_BIND	527	535
FT	CA_BIND	587	595
FT	SITE	1111	1115
FT	DISULFID	71	78
FT	DISULFID	108	126
FT	DISULFID	650	704
FT	DISULFID	768	774
FT	DISULFID	842	858
FT	DISULFID	994	1009
FT	DISULFID	1017	1048
FT	CARBOHYD	33	33
FT	CARBOHYD	86	86
FT	CARBOHYD	185	185
FT	CARBOHYD	646	646
FT	CARBOHYD	667	667
FT	CARBOHYD	723	723
FT	CARBOHYD	859	859
FT	CARBOHYD	894	894
FT	CARBOHYD	929	929
FT	CARBOHYD	1056	1056
FT	CARBOHYD	1067	1067
FT	SEQUENCE	1165 AA;	128725 MW; DABE3A3F1B1463CB CRC64;
SQ	Query Match	26.7%;	Score 1572; DB 1; Length 1165;
	Best Local Similarity	34.6%;	Pred. No. 1.7e-95;
	Matches	409; Conservative	201; Mismatches 465; Indels 108; Gaps 33;
QY	1	FNLDTENAMTQENARG--FGQSVVLQGSVVVVGAPQEIIVAAHQRSGLYQCDISTGSC	58
Db	24	YNLDVRHVQNFSPFLAGRHFGYRVLQV--GNGVVVVGAPSE--GNSMGNLYQCQPTGDCL	79
QY	59	PIRLQVPVEAVNMSIGLSLAATTSPQLLAGCPVTHQTCSENTYVKGICELFGSLRQOP	118
Db	80	PVTLTS--SNYTSKYLGMTLATDPTSDNLLACDPLGSRCTDONIYLSGLCYLIHENLRGPV	137
QY	119	QKFPALRGCPQEDSDIAFLIDGSGSIIPHFRMRKMWSTVMEQKKSKTLFSLMQVSE	178
Db	138	LQGHGPGYCECIKGNVDLVLFLDGSMSLOODEFEKIVDFMKVWKLSSSYQFAAVQFST	197
QY	179	EFRIHFTPEKFNQNNPNSRLKIPITQLLGRTHATGLRKVRRELNIINGARKNAFKILF	238
Db	198	YRTEFTFLDYIROKDPDALLAGVXHMRLTNTFGAINVYVAKEVFRPDLGARPDATKVLI	257


```
QY 239 LLTDGEKFGDPLGVEDVPELD-REGVIRYVIGVDAPRSEKSRQELNTVASKPRDHPV 297
Db 258 IITDGEATDE-----HNDAKDIIRVIIGIKNFKTKESQEAHQFASKEVEFVK 309
QY 298 QINNFEALKTIQNLREKIQAIEGTQTQSSSPHEHMSQSFSAATNSGPLSLTVSGSYD 357
Db 310 ILDTFEKLDLFFELQKKIVIEGTQKDTISFNMELSSSGISADLSEGHVGVAVGAKD 369
QY 358 WAGGVF-LYTSKSKSTFINNTRVDSMDNDAYLGAIA-IILNRVQSLVLGAPRYQHIGL 415
Db 370 WAGGFELDLKADLKSSTVGVNEPLTVESRAGLYGVTVLWPSRGTMSLLATGAPRYQHVGR 429
QY 416 VAMFRO--NFGMESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVLIGAPHYIEQTRGG 473
Db 430 VLLFQDPKRGPSQIQEIDGIGISYFGGELCGVDVDRDGETELLIIAAPLYGQRG 489
QY 474 QVSVCLPRQGRARWQCDVLYGEOGPWGRFGAALTVLGVNDGDKLTVAIGAPGHEEN 533
Db 490 RVFIY---QKIQLFQWVSELQETGYPLGRFGAAIAALTIDNGDELTVDVAGPLEE-- 544
QY 534 RGAVLFGHTSGSGISPSHSQRIAGSLPRQYFQCSLGGQDLTMDGLVDLTVGAQGH 593
Db 545 OGAVYIFNGQOG-GLSPRPSRIEGTQMFSGIQWFGRSHGVKDLGGDLADVAVGAGQ 603
QY 594 VILLRSQPLRVKAIMENPREVARNVFCNDQVVKKEAG-EVRYCLHVQK--STRDRL 650
Db 604 VIVLSRPVVDITTSVSPAPIPVHEVECSYSTSNQKKEGVNLTVCFOVKSLSIT---- 659
QY 651 REGQISQVVTYDALDSGPRHAFVNETKNSRQTVGLTQICETLKLQPCIEDP 710
Db 660 FQGHVLANITYLQLDGHTTRSGGLFPGGKHKLIGNTAVTPV-KSCFVWFHFPIQIDL 718
QY 711 VSPVILRNFSI---VGTPLS--AFGNLRPLVLAEDAQRLLFTALFPFKKNCNDNICQDDL 765
Db 719 ISPINVLSYSLWEIEGTPRDPRALDRDIPPLTKPSPLETKEIFEKNCGDKNCEADL 778
QY 766 SITFSFMSLCLVCGGPREFNVTVVRNDGDSYRTQVTFPPDLDSYRKVSTLQNRSQ 825
Db 779 KLAFSDMRSKILRLTPSASLVRNLRTNDAEYVWQVTLSPQGLSFRKVEIL---KPH 835
QY 826 RSWRLACESASTSVSGALKSTCSINHPIPENSEVTNITFDVDSKASLGNKLLKAN 885
Db 836 SHVPVGCHELEPAVHS-RALSCNVSPIFGEDSMVDIQWENTLQKSGWGFELQAN 894
QY 886 VTS-----ENNPRNTKTEFQLEPLVKYAVVTVTSHGVSKYLNFTASENTRVMQ 937
Db 895 VSCNNESSLEDNSATTS-----IPVYPIVNLTKQENSTLYISFTPKSKIHVVK 947
QY 938 HOYQV---SNLQSRSLPISLVFLVPLRLNTVI---WD---RPQVTFPS-ENLSSTCHTK 986
Db 948 HIYQVRIQPSYDNNP-PLEALVVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESDEA 1006
QY 987 ERLPSSHDFLAELKAPVWNCIAVCQRIQCDIPFGIOEFNATLKNLSFDWYIKTSH 1046
Db 1007 E-----SCSFTG--EFCPIDF---RQELVQVNGVWELRGTIKAS- 1042
QY 1047 NLLIVSTAELFNDSTVFTLLPGCAFVRSQTETKVEFPVNPPLVIGSSVGLLLA 1106
Db 1043 SMLSCLSLAISFNSSKHFLHLYGRNASM-AQVWKVLDLVVEKEMLYLVLSGIGLLLP 1101
QY 1107 LITAAVLKLGFEKQYKDM-----SEGGPPGAPQ 1137
Db 1102 LIFALYKGVFFKRNLERKMEANVDASSEIPGEDAGQPELEKE 1144
```

RESULT 12

```
ITAL HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1170 AA.
AC P20701; O43746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1
```

alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
Name=ITGAL; Synonyms=CD11A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";
RL J. Cell Biol. 108:703-712 (1989).
[2]
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Bichler E.B., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";
RL Genomics 60:295-308 (1999).
[3]
X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036067; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281 (1995).
[4]
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96394682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";
RL Structure 4:931-942 (1996).
[5]
X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition upon luvastatin binding to the CD11a I-domain.";
RL J. Mol. Biol. 292:1-9 (1999).
-!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2, ICAM3 and ICAM4. It is involved in a variety of immune phenomena including leukocyte-endothelial cell interaction, cytotoxic T-cell mediated killing, and antibody dependent killing by granulocytes and monocytes.
-!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L associates with beta-2.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=2;
Name=1;
IsoId=P20701-1; Sequence=Displayed;
Name=2;
IsoId=P20701-2; Sequence=VSP_002738;
Note=No experimental confirmation available;
TISSUE SPECIFICITY: Leukocytes.
-!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins with I-domains do not undergo protease cleavage.
-!- SIMILARITY: Belongs to the integrin alpha chain family.
-!- SIMILARITY: Contains 7 FG-GAP repeats.
-!- SIMILARITY: Contains 1 VWFA domain.
-!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -


```

Db 490 RGRVFYI-----ORRLGFEEVSELOQDGPYPLGRFGEAITALTIDINGDLVDVAVGAP 544
Qy 529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGVLDTV 588
Db 545 LEE--QGAIVFNGRHG-GLSPQSPQRIEGTQVLSGIQWFGRSIHGVKDLGEGDLAVAV 601
Qy 589 GAQGHVLLRSQPLVRKATMEFNPREVARNVFECQV-V-KGKEAGEVRVCLHVVQKSTR 647
Db 602 GAESQMTVLSSRPVVDVMTLMSFSPABIPVHEVECSYSTSNKMEGVNITICFQI-KSLY 660
Qy 648 DRLREGIQSVVTVYDLALDSGRPHSRVAFNETKNSRTOVQLGLTOTCETLKLQLPNCI 707
Db 661 PQP-QGRLVANLTYTLQDGHRTTRRGLFPGRHELNRNIAVT-TSMSCDTFSFHFVQV 718
Qy 708 EDPVSPVILRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQLFTALFPEKNCGN 757
Db 719 QDLISPINVSLNFSLWEEETPRDQRAQKQDIPILRPLSHSEWEI-----PFRNCGE 773
Qy 758 DNICQDLSITTFMSLDCLVVGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVS 817
Db 774 DKCEANLRVSFSPARSRALRLTAFASLSVELSILNEEDAYVQLDLHFPPLGSLFRKVE 833
Qy 818 TLQNSQRSWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKAS 875
Db 834 ML---KPHQIPVSCHELPEESLLRAL---SCNVSSPIFKAGHSVALQMMENTLVNSS 887
Qy 876 LGNKLKLANVTSENN-----MPRTNKTEFOLELPVKYAVVMVVTSHGVSTKYLNFASEN 931
Db 888 WGDSEVLANVTNNEDSDLLDENSATTI---IPILYPINILIQDQEDSTLYVSFTPKGP 944
Qy 932 TSRVMQHOYOV---SNLQORSLP-ISLVLVPLVRLNQTVMIDRPQVTFSENLSSTCHK- 986
Db 945 KIHQVKHMYQVARIQPSIHDPNIPTLEAVVGVPOPPSEGPITHQWSVQMEPPV--PCHYED 1002
Qy 987 -ERLPSHSD--FLAELRKAPVNCSTAVCQRQCQDIPFGIQEEFNATLKGMLSPDWYTK 1043
Db 1003 LERLPDAEPCPLGALFRCPV-----FRQELVQVIGTLELVGEYE 1044
Qy 1044 TSHNHLIVSTAEILLFNDSVFTLLPQGAFAVSQTETKVEPEVFNPLPIVSSVGGILL 1103
Db 1045 AS-SMFLSCSLISISFNSSKPHLYGNSASL-AQVVMKVDVYVQKMLYLVLSGIGILL 1102
Qy 1104 LLALITAAALYKLGFFKRYQKMMSEG-GPPGAEP 1136
Db 1103 LLLLIIFVLYKVGFFKRLKKEKAGRGVNGIP 1136

RESULT 13
Q6TYB8
ID Q6TYB8 PRELIMINARY; PRT; 1166 AA.
AC Q6TYB8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
OS Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OC Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumrikat P., Kannan M.S., Maheswaran S.K.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL: AY382558; AAQ9005.1; -.
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 1.
DR Pfam: PF00357; Integrin_alpha; 1.
DR Pfam: PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.

```

```

DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02034; VWF; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8DBAFBA896C9DF CRC64;

Query Match
Best Local Similarity 34.3%; Score 1546; DB 2; Length 1166;
Matches 405; Conservative 202; Mismatches 469; Indels 106; Gaps 33;

Qy 1 FNDLTENAMTFQENARG--FGQSVVQLQGSRRVVVGAPOEIVAAQNRGSLQCDYSTGCE 58
Db 25 YNLDVRHVQNFSPFLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLQCCPETGCDL 80
Qy 59 PIRLQVPVEAVNMSLGLSLAATSPQLLACGTVHTQTCSENTYVVKGLCFLFGSNLRQP 118
Db 81 PVTLS--SNYTSKYLGMTLATDFTSNLLACDPLGSLRTCDQNTLYSLGLYLHLENLRGPV 138
Qy 119 QKPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEWVSTVMEQLKSKTLFLSLWOYSE 178
Db 139 LQGHGPGYQECIKGNVDLVFLFDGMSLQODEFEKIVDFMKDVMKLSNSSYQAAVQFST 198
Qy 179 ERIHFTFEFQNNPNRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILF 238
Db 199 YFRTTEFTFDYIKQKDPDALLAGVKHMLLTNTFGAINVVAKEVFRPDILGARDATKVL 258
Qy 239 LLTDEKFGDPLGYEDVIPELDREGVIRVIGVGAFAFRSEKROELNVTVASKPRDHVQ 298
Db 259 IITDGK---PPNTTILMRPKTSS---RSLLGIGNKFKTESQEAHLQFAPVEEFVKI 311
Qy 299 INNFALKTIQNLREKIFAIEGTQTGSSSEFHEMSOEGFSAATISNGPLSTVGSYDW 358
Db 312 LDTEFKLDLFTLQKIVVIGTSKQDLTSFNMELSSGSIADSEGHVGVGAVGAKDW 371
Qy 359 AGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGA---IILNRVQSLVGLAPRYOHIGIV 416
Db 372 AGGFDLKLADLKSSTFVGNELPTVESRAGLYGTVTRLPSTRGTMSSLLATGAPKYOHVGRV 431
Qy 417 AMFRO--NTGMWESNANVGTQIGAYFGASCLSDVDVDSNGSTDVLVIGAPHYEYTRGO 474
Db 432 LLFQPKRGGPWSQIEIDGIGISYFGGELCGVDVDRDGETELLLIAAPLYYGRQGR 491
Qy 475 VSVCPPLRGQARWQCDALYCEQOPWGRFGAALTGLDVNGDKLTDTVAIGAPCEENR 534
Db 492 VFIY---QKIQLEFQWSELOGETYPLGRFGAATAALTIDNGDELTVAVGAPLEE--Q 546
Qy 535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVLDITVGAQGHV 594
Db 547 GAVYIFNGQOG-GLSPRPSQRIEGTQMPSGIQWFGRSIHGVKDLGGDLADVAVGAEQV 605
Qy 595 LLRSQPLVRKATMEFNPREVARNVFECQVVKGEAG-EVRVCLHVQK--STRDLRL 651
Db 606 IVLSSRPVDIITSVSPAPIPVHEVECSYSTSNKKEGVNLTVCQVKSLSIST---F 661
Qy 652 EQIQSVVTVYDLALDSGRPHSRVAFNETKNSRTOVQLGLTOTCETLKLQLPNCIDPV 711
Db 662 QGHLVANLTYTLQDGHRTTRRGLFPGRHELNRNIAVTTPV-KSCFVFWFHFICIDLI 720
Qy 712 SPIVLRNFSI---VGTPLS--AFGNLRPVLAEDAQLFTALFPEKNCGNINDICQDLS 766
Db 721 SPINVSLSYSLWEEETPRDRLDRDIPILKPSHLETKEIPFEKNCEDKNCADLK 780
Qy 767 ITFSFMSLDCLVVGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSQR 826
Db 781 LAESDMRSKILRLTPPSASLSVRLTNTAEDAYVQVLTSLFPQGLSFRKVEIL---KPHS 837
Qy 827 SWRLACESASTEVSGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANV 886
Db 838 HVPVGCHELPEEAVVHS-RALSNCVSSPIFGESWDIQQWENTLQKSGWDFIELQANV 896
Qy 887 TS-----ENNPRNTKTEFOLELPVKYAVVMVVTSHGVSTKYLNFNTASNTSRVMQH 938

```

```

897 SCNNEDSLLEDNSATTS-----IPVMPINVLTKQENSTLYISTPKSPRIHHVKH 949
939 QYQV-----SNLQORSLPTSLVFLVRLNQTIV---WD---RPOVTF--ENLSSCTHTKE 987
950 IQVRIQPSNYDNMP-PLEALVRVRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008
988 RLPSSHDFLAELKAPVNVNCSIAVCORIQCIDIPFGIQEENFATLKNLSFDWYIKTSHN 1047
1009 -----SCSFGT--EPRCFIDF---ROEILVQVNGMVELRGTIKAS-S 1044
1048 HLLIVSTAEILFENDSVFTLLPGQAFVRSQTEKVEPPEVNPPLVIGSSVGGILLALL 1107
1045 MSLCSSLAISFNSSKHFLHGRNASM-AQVVMKVDLVYEKEMLYLVLSGGILLALLFL 1103
1108 ITAALYKLGFFKQYKQDM-----SEGGPPGABPQ 1137
1104 IFIALLYKVGFFKRLKKEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 14
AAQ90015
ID AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (T-EMBLrel. 27, Created)
DT 02-MAR-2004 (T-EMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (T-EMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a."
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY382558; AAQ90015.1;
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.3%; Score 1546; DB 2; Length 1166;
Best Local Similarity 34.3%; Pred. No. 9.1e-94;
Matches 405; Conservative 202; Mismatches 469; Indels 106; Gaps 33;

QY 1 FNLDENAMTFQENARG--FGQSVVQLQGSRRVVVGAPOEIVAAQNRSLYQCYSTGSC 58
DB 25 YNLDRHVQNFSPFPLAGRHFGYRVLVQV-GNGVVVGAPSE---GNSMGNLYQCQPETG 80
QY 59 PIRLOVPEAVNMISGLSLAATTSPQLLACGPTVHQCSTENTYVKGICFLFGSLNRQ 118
DB 81 PVTLS--SNYTSKYLGMTLATDPTSDNLLACDPLGSLRTCDQNTYLSGLCYLIHENLR 138
QY 119 QKFEALRGCPQEDSIAFLDGGSIIPHFRMKKEVSTVMEQLKKSKTLFSLMOYSE 178
DB 139 LQGHFGYQECIKGNVDLVFLFGMSLQDDEFEKIVDFMKDVMKLSNSYQFAVQFST 198
QY 179 EFRHFTKEFQNNPNRSLKIPITQLLGRTHATGLRKVRVRELPNITNGARKNAFKILF 238
DB 199 YFRTEFTLDYIKQKDPDALLAGVXKMRLLTNTFGAINYVAKEVFRPDLGARPDATK 258
QY 239 LLTDEKFGDPLGYEDVPELDRGVIRYVIGVGDAFRSEKRSQELNIVASKPPRDHVF 298
DB 259 IITDGK---PPTNTILMRPKTS-----RSLLGIGNKFKESQEAHLHOFASKPVFE 311
QY 299 INNFALKTIQNLREKIPAEIGTQTGSSSSSEHEMSQEGFAAITSNGPLLSYGVSD 358
DB 312 LTFEKKDLFLTELQKKIVIEGTSKQDLTSFNWELSSSGISADLSEGHVVGAVGAKDW 371
QY 359 AGGVF-LYTSKEKSTFINMTRVDSQNDAYLGYAAA-IILNRVQSLVLGAPYOHIGLV 416
DB 372 AGGFLLDLKLSSTFVGNELPTVESRAGYLGTVYTRLPSTRGMTSLLATGAPKYOHVGR 431

```

```

417 AMFRQ--NTGMWSENANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLICAPHYVEQTRGO 474
432 LLFQPKRGGPWSQIEIDGIIQISYFGGELCGVDVDRDGETELLIIAALPYIGQORGR 491
475 VSVCPPLPGQARWQCDVLYCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534
492 VFYI---KQILQFQMVSELOQETGYPLGRFGAALTAALTDINGDELTDVAVGAPLEE--Q 546
535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVAQGHV 594
547 GAVYIFNGQOG-GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDLGADVAVGAEQV 605
595 LLLRSQPLRVKAIMFEMFPRVARNVFECNDQVVKGEAG-EVRVCLHVQK--STRDLR 651
606 IVLSSRPVVDIITISVSFPAEIPVHEVECSYSTSNQKKEGVNLTVCQVKSILIST----F 661
652 EQIQSVVYTDALDPSGRHSRAVENETKNSTRROTQVLGLTQTCTELKLQLPNCIEDPV 711
662 OGHLVANLYTTLQDGHRTSRGLPFGGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720
712 SPIVLRLNPSL---VGTPLS--AFGNLRPLVLAEDAQRLLFTALFPFKKNGCNDNICOD 766
721 SPINVSLSYLWEEEGTDPDRALDRDIPILKPSPHLETKEIPFEKNCGEDKNCEADLK 780
767 ITFSFMSLDCLVVGPREFNVTVTVNDGEDSVYRTQVTFPFDLDSVRKVSTLQNRSOR 826
781 LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYVWQVTLSPFOGLSFRKVEILL---K 837
827 SWRLACESASTVSGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANV 886
838 HVPVGCBELEPEEAVVHS-RALSCNVSPFGEDESMVDIQWVFNTLQKSGWDFELQANV 896
887 TS-----ENNMPRTNTEFQLELPVKYAVVMTVSHGVSTKYLNFTASENTSRVNOH 938
897 SCNNEDSLLEDNSATTS-----IPVMPINVLTKQENSTLYISTPKSPRIHHVKH 949
939 QYQV-----SNLQORSLPTSLVFLVRLNQTIV---WD---RPOVTF--ENLSSCTHTKE 987
950 IQVRIQPSNYDNMP-PLEALVRVRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1008
988 RLPSSHDFLAELKAPVNVNCSIAVCORIQCIDIPFGIQEENFATLKNLSFDWYIKTSHN 1047
1009 -----SCSFGT--EPRCFIDF---ROEILVQVNGMVELRGTIKAS-S 1044
1048 HLLIVSTAEILFENDSVFTLLPGQAFVRSQTEKVEPPEVNPPLVIGSSVGGILLALL 1107
1045 MSLCSSLAISFNSSKHFLHGRNASM-AQVVMKVDLVYEKEMLYLVLSGGILLALLFL 1103
1108 ITAALYKLGFFKQYKQDM-----SEGGPPGABPQ 1137
1104 IFIALLYKVGFFKRLKKEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 15
ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=Itgal; Synonyms=Lfa-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1

```


us-09-902-481b-4.rup

Tue Nov 9 12:56:08 2004

Db 721 SPINVSINFLSLEEBEGTPRQKRAMQPIILRPSIHV-TKEIPFEKNCGEDKKCEANLTL 779
QY 768 TFSFMSLDCLVVGSP-----REFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQN 821
Db 780 SSPARS-----GFLRLMSASLAVETLSNGEDAYWRLDLDPRGLSFRKVMQLQ- 831
QY 822 QRSQRSWRLACESASSTEVS GAL-KSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKL 880
Db 832 --PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQVMFNTLNSSWEDFV 887
QY 881 LKANVTSEN-NMPRTNKTFOLELPVKYAVVMVTVSHGVSTKYLNFASENTSRVMOHQ 939
Db 888 ELNGTVHCENENSSLQEDNSAATHIPVLPVNLTKBOENSTLYISFTPKGPKTOQVQHV 947
QY 940 YQVSNLQORSPLISLVFLVPRLNQTIVWDRPQ-----VTFSENLS----TCHTKE-RLP 990
Db 948 YQV-----RIQFSAYDHNMT-LEALVGVPRHSEDLITYTWSVQTDPLVTCHSEDLKRP 1001
QY 991 SHSDFLAELRKAPVWNCISIAVCORIOCDIPFGIOBEFNATLKNLSFDWYIKTSHNHL 1050
Db 1002 SSE---AEQPCLPV-----QFRCPVIF---RWEILIOVTGTVELSKEIKAS-STLS 1046
QY 1051 IVSTABILFNDVSFTLLPGQAFVRSQTEKVEPPEVNPPLPLIVGSSVGGILLALITA 1110
Db 1047 LCSSLVSFNSSKHFLYGSKA-SEAQVLVKVDLIHEKEMLVHVVLSGIGLVLLFLFL 1105
QY 1111 ALYKLGFFKQYKDM-SEGPPGAEP 1136
Db 1106 ALYKVGFFRNLKKEADGGVNGSP 1132

Search completed: November 9, 2004, 12:17:33
Job time : 142.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5852	99.5	1153	2	Aaw65090 Human Bet
2	5852	99.5	1153	3	Aab07360 Human CDI
3	5852	99.5	1153	5	Aau80252 Human int
4	5852	99.5	1153	5	Abg61469 Human Bet
5	5852	99.5	1153	5	Abg61469 Human Bet
6	5852	99.5	1153	7	Aa014428 Integrin
7	5842	99.3	1153	7	Add25615 Binding d
8	5836.5	99.2	1152	8	Aar04136 Alpha sub
9	5829.5	99.1	1152	8	Adm99589 Human int
10	3473	59.0	1163	8	Adp12435 Protein e
11	3459	58.8	1163	8	Adp44061 Human CDI
12	3459	58.8	1163	8	Adn02004 Human inf
13	3450	58.6	1163	8	Adq17510 Human sof
14	3436	58.4	1163	2	Aar07120 p150.95 a
15	3436	58.4	1163	2	Aaw65091 Human Bet
16	3436	58.4	1163	3	Aab07361 Human CDI
17	3434	58.4	1163	6	Abg61470 Human Bet
18	3434	58.4	1163	7	Abu07406 Protein d
19	3401	57.8	1161	7	Adg32005 Human hom
20	3401	57.8	1161	2	Aar78166 Human bet
21	3401	57.8	1161	2	Aaw23049 Human bet
22	3401	57.8	1161	2	Aaw57491 Human bet
23	3401	57.8	1161	2	Aaw65089 Human bet
24	3401	57.8	1161	2	Aaw72825 Human alp
25	3401	57.8	1161	3	Aaw73342 Human alp
					Aab07359 Human alp

26	3401	57.8	1161	5	Abg61468	Human Bet
27	3385.5	57.5	1161	2	Aaw23064	Human Bet
28	3385.5	57.5	1161	2	Aaw65106	Human Bet
29	3385.5	57.5	1161	2	Aaw72837	Human alp
30	3385.5	57.5	1161	2	Aaw73343	Human alp
31	3385.5	57.5	1161	3	AAB07376	Human alp
32	3385.5	57.5	1161	5	ABG61485	Human Bet
33	3226.5	54.8	1161	2	AAR78169	Human Bet
34	3224.5	54.8	1161	2	Aaw23062	Rat beta
35	3224.5	54.8	1161	2	Aaw60004	Rat alpha
36	3224.5	54.8	1161	2	Aaw72824	Rat alpha
37	3224.5	54.8	1161	3	AAB07374	Rat alpha
38	3224.5	54.8	1161	3	ABG61483	Rat beta2
39	3217.5	54.7	1161	2	ABG61483	Rat beta2
40	3217.5	54.7	1161	2	Aaw73345	Mouse alp
41	3212	54.6	1161	2	Aaw23061	Mouse alp
42	3212	54.6	1161	2	Aaw60003	Mouse alp
43	3212	54.6	1161	2	Aaw65103	Mouse alp
44	3212	54.6	1161	2	Aaw72836	Mouse alp
45	3212	54.6	1161	2	Aaw73347	Mouse alp

ALIGNMENTS

RESULT 1

Aaw65090
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or

XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is
XX used to describe a method for identifying compounds that modulate the
XX interaction of the beta-integrin alpha-d subunit with a binding partner
XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX -d binding partner, one of which is immobilised and the other of which is
XX labelled, in the presence of a test compound, and determining if the
XX compound affects binding between the alpha-d polypeptide and alpha-d
XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX comprising the cytoplasmic, transmembrane or extracellular domain of
XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ	Sequence 1153 AA;	
Query Match	99.5%; Score 5852; DB 2; Length 1153;	
Best Local Similarity	99.2%; Pred. No. 0;	
Matches 1128; Conservative	7; Mismatches 2; Indels 0; Gaps 0;	
QY	1 FNLDTENAMTFQENARFGQSVVQLOGSRVVGAPQEIIVAANQKSLYQCDYSTGSCPEI 60	
DB	17 FNLDTENAMTFQENARFGQSVVQLOGSRVVGAPQEIIVAANQKSLYQCDYSTGSCPEI 76	
QY	61 RLQVPEAVNMSLGLSLAATTPPQLAACQPTVHQTCTSENTYVKGLCFLFGSNLRQOPQK 120	
DB	77 RLQVPEAVNMSLGLSLAATTPPQLAACQPTVHQTCTSENTYVKGLCFLFGSNLRQOPQK 136	
QY	121 FPEALRCQPOEDSDIAFLIDGSGSIIIPDPRMKEWSTVMEQLKSKTFLFSLMOYSEEF 180	
DB	137 FPEALRCQPOEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQKKSKTFLFSLMOYSEEF 196	
QY	181 RIHFTFEFQNNPNSLSIKETITQLGRTHATGLRKVVRLEFNTNGARKNAFKILFL 240	
DB	197 RIHFTFEFQNNPNSLSIKETITQLGRTHATGLRKVVRLEFNTNGARKNAFKILV 256	
QY	241 TDGEKPGDPLGYEDVPELDBREGVIRVVGDAFRSEKSHOELNTVASKRPHRVQIN 300	
DB	257 TDGEKPGDPLGYEDVPEADREGVIRVVGDAFRSEKSHOELNTVASKRPHRVQIN 316	
QY	301 NFEALKTIONQLREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAG 360	
DB	317 NFEALKTIONQLREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGELLSTVGSYDWAG 376	
QY	361 GVFLYTSKEKSTFTNMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420	
DB	377 GVFLYTSKEKSTFTNMTVDSDMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436	
QY	421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTGGQSVCP 480	
DB	437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTGGQSVCP 496	
QY	481 PRGQARWCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540	
DB	497 PRGQARWCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 556	
QY	541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLRLRSQ 600	
DB	557 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLRLRSQ 616	
QY	601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660	
DB	617 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676	
QY	661 YDLALDSGRPHSRVAFNETKNSRTQVGLGTQTCETILKLPNCIEDPSPVILRLNF 720	
DB	677 YDLALDSGRPHSRVAFNETKNSRTQVGLGTQTCETILKLPNCIEDPSPVILRLNF 736	
QY	721 SLVGTPLSAFAGNLRPVLAEDAQRLETFALFPPEKNGNDNICQDDLSTIFSFMSLCLVVG 780	
DB	737 SLVGTPLSAFAGNLRPVLAEDAQRLETFALFPPEKNGNDNICQDDLSTIFSFMSLCLVVG 796	
QY	781 GPREFNVTVNRDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWELACESASTEV 840	
DB	797 GPREFNVTVNRDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWELACESASTEV 856	
QY	841 SCALKSTSCSINHPIFENSEVTFNTTFVDKASLGNKLLKANVTSENNNPRNKTFEF 900	
DB	857 SCALKSTSCSINHPIFENSEVTFNTTFVDKASLGNKLLKANVTSENNNPRNKTFEF 916	
QY	901 QLELPVKYAVYVMVVTSHGVSTKYNFTASENTSRVVMQHOYQVSNLQORSLPSLFLVPV 960	
DB	917 QLELPVKYAVYVMVVTSHGVSTKYNFTASENTSRVVMQHOYQVSNLQORSLPSLFLVPV 976	

QY	961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCIAVCQRIQCDIP 1020	
DB	977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCIAVCQRIQCDIP 1036	
QY	1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET 1080	
DB	1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQCAFVRSQTET 1096	
QY	1081 KVPEFPEVNDPLPIVGVSSVGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137	
DB	1097 KVPEFPEVNDPLPIVGVSSVGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153	
RESULT 2		
AAB07360	standard; protein; 1153 AA.	
ID	AAB07360	
XX		
AC	AAB07360;	
XX		
DT	17-JAN-2001 (first entry)	
XX		
DE	Human CD11b protein sequence.	
XX		
KW	Human; macrophage infiltration inhibition; alpha_d integrin;	
KW	leukocyte integrin; Leu-CAM; leukointegrin; immune response;	
KW	inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;	
KW	atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;	
KW	lung inflammation; acute respiratory distress syndrome; Crohn's disease;	
KW	rheumatoid arthritis; central nervous system injury; CD11b.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200029446-A1.	
XX		
PD	25-MAY-2000.	
XX		
PF	16-NOV-1999; 99WO-US027139.	
XX		
PR	16-NOV-1998; 98US-00193043.	
PR	08-JUL-1999; 99US-00350259.	
XX		
PA	(ICOS-) ICOS CORP.	
XX		
PI	Gallatin MW, Van Der Vieren M;	
XX		
DR	WPI; 2000-387751/33.	
XX		
PT	Use of novel anti-alpha integrin d monoclonal antibodies to inhibit	
PT	macrophage infiltration and reduce inflammation at central nervous system	
PT	injury sites.	
XX		
PS	Example 5; Fig 1; 270pp; English.	
XX		
CC	Integrins are a class of membrane-associated molecules that participate	
CC	in cellular adhesion. Integrins are made up of an alpha subunit and a	
CC	beta subunit. One class of human integrins are restricted to expression	
CC	in white blood cells and have a common beta2 subunit: the leukocyte	
CC	integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins	
CC	have an important role in immune and inflammatory responses. The present	
CC	protein sequence is the human integrin alpha subunit CD11b. This sequence	
CC	was used in an alignment to identify a novel beta2 integrin alpha	
CC	subunit: alpha d (AAA60014 and AAB07359). The present sequence has	
CC	approximately 60% identity to the protein sequence of alpha d. The	
CC	Alpha d gene and protein may be useful in therapy for diseases linked to	
CC	aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple	
CC	sclerosis, asthma, psoriasis, lung inflammation, acute respiratory	
CC	distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency	
CC	(IAD). In addition, anti-alpha d monoclonal antibodies may be used in the	
CC	inhibition of macrophage infiltration at the site of a central nervous	
CC	system injury. The monoclonal antibodies can also be used to detect and	
XX	diagnose Crohn's disease	

[illegible]

XX	Sequence 1153 AA;	Query Match	Best Local Similarity	Matches 1128; Conservative	99.5%; Score 5852; DB 5; Length 1153;	7; Mismatches	2; Indels	0; Gaps
QY	1 FNLDENAMTFOENARGFGQSVQLOQSRVVVGAPQEIIVANQSGSLVQCDYSTGSCPEI 60							
DB	17 FNLDENAMTFOENARGFGQSVQLOQSRVVVGAPQEIIVANQSGSLVQCDYSTGSCPEI 76							
QY	61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPQK 120							
DB	77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPQK 136							
QY	121 FPEALRGCPQSDSDIAFLIDGSGSIIIPDPRFMKEWSTVMEQLKSKTILSLMOYSEEF 180							
DB	137 FPEALRGCPQSDSDIAFLIDGSGSIIIPDPRFMKEFVSTVMEQLKSKTILSLMOYSEEF 196							
QY	181 RIHFTFKFQNNPNSRLIKPITOLLGTHTHATGLRKVVRELFNITNGARKNAKILFL 240							
DB	197 RIHFTFKFQNNPNSRLIKPITOLLGTHTHATGLRKVVRELFNITNGARKNAKILVVI 256							
QY	241 TDGEKFGDPLGYEDVIPBLDREGVIRYVIGDAPFRSEKSRQELNTVASKPPRDHVFQIN 300							
DB	257 TDGEKFGDPLGYEDVIPADREGVIRYVIGDAPFRSEKSRQELNTIASKPPRDHVFQVN 316							
QY	301 NFEALKTIQNLREKIPFAIEGHTQTCSSSPFHEMSQEGFSAATISNGPLLSTVGSYDWAG 360							
DB	317 NFEALKTIQNLREKIPFAIEGHTQTCSSSPFHEMSQEGFSAATISNGPLLSTVGSYDWAG 376							
QY	361 GVFLTSKEKSTFINNTRVDSMDNAYLGAAAIILNRNVSQVLGAPRYOHIGLVAMFR 420							
DB	377 GVFLTSKEKSTFINNTRVDSMDNAYLGAAAIILNRNVSQVLGAPRYOHIGLVAMFR 436							
QY	421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480							
DB	437 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 496							
QY	481 PRGQARWQCDVAVLGEQGPWGRFGAALTVLGVNGDKLTDVAIGAFGEEDNRGAVYLF 540							
DB	497 PRGQARWQCDVAVLGEQGPWGRFGAALTVLGVNGDKLTDVAIGAFGEEDNRGAVYLF 556							
QY	541 HGTSGSGISPSHSORISAGSKLSPRLOQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600							
DB	557 HGTSGSGISPSHSORISAGSKLSPRLOQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616							
QY	601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSQVVT 660							
DB	617 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQTSQVVT 676							
QY	661 YDLALDSGRPHSRAVFNETKNSRQTQVLGHTQTCETLKLQPLNCIEDPVSPIVRLNF 720							
DB	677 YDLALDSGRPHSRAVFNETKNSRQTQVLGHTQTCETLKLQPLNCIEDPVSPIVRLNF 736							
QY	721 SLVGTPLSAFAGNLRPVLAEDARLFTALPPPEKNCNDNICODDISITFSEMSLCLVVG 780							
DB	737 SLVGTPLSAFAGNLRPVLAEDARLFTALPPPEKNCNDNICODDISITFSEMSLCLVVG 796							
QY	781 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLRYKVTSTLQNSQRSMFLACESASSTEV 840							
DB	797 GPREFNVTVVRNDGSDSYRTQVTFPPFLDLRYKVTSTLQNSQRSMFLACESASSTEV 856							
QY	841 SGALKSTCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRNKTKEF 900							
DB	857 SGALKSTCSINHPIIPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRNKTKEF 916							
QY	901 QLELPVKYAVYMWVTSHGVSSTKYNLTASNTSRVMQHOYQVSNLQORSLPLSLVFLVPV 960							
DB	917 QLELPVKYAVYMWVTSHGVSSTKYNLTASNTSRVMQHOYQVSNLQORSLPLSLVFLVPV 976							
QY	961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP 1020							

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 5; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI	76
QY	61	RLOVPVEAVNMSLGLSLAATSPOLLACGTVHTQTCSENTYVKGCLCFLFSGNLRQQPQK	120
DB	77	RLOVPVEAVNMSLGLSLAATSPOLLACGTVHTQTCSENTYVKGCLCFLFSGNLRQQPQK	136
QY	121	FPEALRGCPQSDSDIAFLIDSGSIIIPHDRFRMEKWTVMQOLKSKTFLSLMOYSEEP	180
DB	137	FPEALRGCPQSDSDIAFLIDSGSIIIPHDRFRMEKWTVMQOLKSKTFLSLMOYSEEP	196
QY	181	RIHFTFEFQNNPNRSLIKPITQLLGRTHPTATGLRKVVRELFNITNGARKNAFKILFL	240
DB	197	RIHFTFEFQNNPNRSLIKPITQLLGRTHPTATGLRKVVRELFNITNGARKNAFKILV	256
QY	241	TDGKFKGDPGLGYEDVIBELDEGVIRYVIGVDAFRSEKSRQELNITVASKPPRDHVFQIN	300
DB	257	TDGKFKGDPGLGYEDVIBELDEGVIRYVIGVDAFRSEKSRQELNITVASKPPRDHVFQIN	316
QY	301	NFEALKTIQNLREKIFAIEGTQTGSSSPFEHMSQEGFSAATITSGPLLSITVGSYDWAG	360
DB	317	NFEALKTIQNLREKIFAIEGTQTGSSSPFEHMSQEGFSAATITSGPLLSITVGSYDWAG	376
QY	361	GVFLYTSKESKSTFINMTRVDSMDNDAYLVAAAIILNRVQSLVIGAPRYQHIGLVAMFR	420
DB	377	GVFLYTSKESKSTFINMTRVDSMDNDAYLVAAAIILNRVQSLVIGAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	480
DB	437	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	496
QY	481	PRGQARWQCDVLYGEOGQWGRFGAALTVDGVNDGKLTVAIGAPGEDNRGAVYLF	540
DB	497	PRGQARWQCDVLYGEOGQWGRFGAALTVDGVNDGKLTVAIGAPGEDNRGAVYLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ	616
QY	601	PVLVRKAIEMFNEPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT	660
DB	617	PVLVRKAIEMFNEPREVARNVFECDQVVKKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT	676
QY	661	YDLALDSGRPHSAVFNENKSTRROTQVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720
DB	677	YDLALDSGRPHSAVFNENKSTRROTQVGLGLTQTCETLKLQLPNCIEDPVPVILRLNF	736
QY	721	SLVGTPLSAFNLRLPVLAEQAORLFTALPFPEKNCNDNIQDDLSITFSFMSLDCLVWG	780
DB	737	SLVGTPLSAFNLRLPVLAEQAORLFTALPFPEKNCNDNIQDDLSITFSFMSLDCLVWG	796
QY	781	GPREFNVTVVRNDGDSYRTQVTFPPDLDSVRKYSTIQRQSRQSWRIACESASSTEV	840
DB	797	GPREFNVTVVRNDGDSYRTQVTFPPDLDSVRKYSTIQRQSRQSWRIACESASSTEV	856
QY	841	SGALKSTCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	900
DB	857	SGALKSTCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	916
QY	901	QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSILPSLVLFPV	960
DB	917	QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVVMQHQYQVSNLQORSILPSLVLFPV	976
QY	961	RLNQTWIDRPQVTFSENLSSTCHTYERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP	1020
DB	977	RLNQTWIDRPQVTFSENLSSTCHTYERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP	1036

QY	1021	PFQIOBEFNATLKNLSFDWIKTSHNHLIIIVSTABILFNDSVFTLLPGQAFVRSQTE	1080
DB	1037	PFQIOBEFNATLKNLSFDWIKTSHNHLIIIVSTABILFNDSVFTLLPGQAFVRSQTE	1096
QY	1081	KVEFEFVNPPLPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMMSGGPPGABPQ	1137
DB	1097	KVEFEFVNPPLPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMMSGGPPGABPQ	1153
RESULT	5		
AAO14428			
ID	AAO14428	standard; protein; 1153 AA.	
XX	AAO14428;		
AC	AAO14428;		
XX	03-MAY-2002	(first entry)	
DT			
XX			
DE		Integrin Mac-1 alpha subunit.	
XX			
KW		Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
KW		open conformation; integrin related inflammatory disorder;	
KW		integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW		reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW		viral infection; cancer; gene therapy; vaccine;	
KW		bioactive agent screening.	
XX			
OS		Unidentified.	
XX			
PN	WO200204521-A2.		
PD	17-JAN-2002.		
XX			
PF	09-JUL-2001; 2001WO-US021805.		
XX			
PR	07-JUL-2000; 2000US-0216600P.		
XX			
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.		
PA	(BLOO-) CENT BLOOD RES.		
XX			
PI	Springer T;		
XX			
DR	WPI; 2002-148167/19.		
XX			
PT	New integrin I domain protein having alteration in at least 2		
PT	noncontiguous regions and exits in an open conformation, useful for		
PT	treating, preventing or suppressing inflammatory or immunological		
PT	disorders.		
XX			
PS	Example 1; Fig 1F; 90pp; English.		
XX			
CC	The invention comprises structurally biased variant integrin inserted (I)		
CC	domain proteins, wherein the alterations to the protein occur in at least		
CC	two noncontiguous regions. Specifically the variant integrin I domain		
CC	proteins are structurally biased to exist in the open conformation,		
CC	thereby altering the binding ability of the protein. The invention also		
CC	comprises nucleic acids encoding the variant integrin I domain proteins.		
CC	The integrin I domain proteins and nucleic acids are useful for treating,		
CC	preventing or suppressing integrin related inflammatory and immunological		
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain		
CC	proteins and nucleic acids can also be used for treating: ischaemia/		
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral		
CC	infection; and cancer. The variant integrin I domain nucleic acids and		
CC	proteins may be used in gene therapy, as vaccines and to screen for		
CC	bioactive agents. The present amino acid sequence represents the Mac-1		
CC	alpha subunit of integrin		
XX			
SQ	Sequence 1153 AA;		

Query Match 99.5%; Score 5852; DB 5; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFQENARGFGQSVVOLQGSRRVVVGAPQEIIVAAVORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVOLQGSRRVVVGAPQEIIVAAVORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAAATTPPOLACGPTVHQTGSENTYVKGCLFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAAATTPPOLACGPTVHQTGSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRCPOEDSDIAFLIDGSGSIIIPHDPRMKEMVSTVMEQKKKTLFSLMQYSEEF 180
Db 137 FPEALRCPOEDSDIAFLIDGSGSIIIPHDPRMKEMVSTVMEQKKKTLFSLMQYSEEF 196
QY 181 RIHTTFEQQNNPRSLRKPITQILGRTHTATGLRKVRELFNITGARKNAFKILFL 240
Db 197 RIHTTFEQQNNPRSLRKPITQILGRTHTATGLRKVRELFNITGARKNAFKILFL 256
QY 241 TDGKFGDPLGYEDVPELDEGVIRYVIGVDGAFRSEKSKQELNTVASKPRPHVFQIN 300
Db 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSKQELNTIASKPRPHVFQVN 316
QY 301 NFEALKTIONQIREKIFAIBGTQTGSSSSPHEHMSQEGFSAATISNGPLLSVGSYDWAG 360
Db 317 NFEALKTIONQIREKIFAIBGTQTGSSSSPHEHMSQEGFSAATISNGPLLSVGSYDWAG 376
QY 361 GFELYTSKEKSTFINMTREVDSMDNDAYLGYAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GFELYTSKEKSTFINMTREVDSMDNDAYLGYAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLICAPHYIEQTRGGQVSVCP 480
Db 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLICAPHYIEQTRGGQVSVCP 496
QY 481 PRGRARQWQDAVLGEGQPMWRFGAALTVLGDVNGDKLTDVAIGABGEDNRCGAVILF 540
Db 497 PRGRARQWQDAVLGEGQPMWRFGAALTVLGDVNGDKLTDVAIGABGEDNRCGAVILF 556
QY 541 HGTSGSGLSPSHSQRISAGSKLSPRLOYFGQSLGGQDLTMDGLVDLTVGAQHVLLLSRQ 600
Db 557 HGTSGSGLSPSHSQRISAGSKLSPRLOYFGQSLGGQDLTMDGLVDLTVGAQHVLLLSRQ 616
QY 601 PVLVKALMEPNPREVAENVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVKALMEPNPREVAENVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNETKSTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
Db 677 YDLALDGRPHSRVAFNETKSTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVILRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFEKNGCNDNICODLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADQRLFTALPPFEKNGCNDNICODLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWELACESASTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWELACESASTEV 856
QY 841 SGALKSTCSINHPITPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEP 900
Db 857 SGALKSTCSINHPITPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEP 916
QY 901 QLELPVKAVYVNVVTSYKVLNFTASENTSRVMOHQYQVSNLQORSLPISLFLVFPV 960
Db 917 QLELPVKAVYVNVVTSYKVLNFTASENTSRVMOHQYQVSNLQORSLPISLFLVFPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAEIRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAEIRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVFTLLPQGGAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSHNLLIVSTAILFNDSVFTLLPQGGAFVRSQTET 1096
QY 1081 KVPEPFEVNPPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1137
Db 1097 KVPEPFEVNPPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1153

RESULT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX AC ADD25615;
XX DT 15-JAN-2004 (first entry)
XX DE Binding domain-immunoglobulin fusion protein-associated protein #85.
XX KW Binding domain; immunoglobulin; fusion protein; cytostatic;
XX KW antiarthritic; immunosuppressive; antidiabetic; antichyroid;
XX KW neuroprotective; hinge region; immunoglobulin heavy chain;
XX KW CH2 constant region; CH3 constant region; IgG1;
XX KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
XX KW malignant condition; B-cell disorder; melanoma; sarcoma;
XX KW rheumatoid arthritis; myasthenia gravis; Grave's disease;
XX KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX OS Unidentified.
XX PN US2003118592-A1.
XX PD 26-JUN-2003.
XX PF 25-JUL-2002; 2002US-00207655.
XX PR 17-JAN-2001; 2001US-0367358P.
XX PR 17-JAN-2002; 2002US-00053530.
XX PR 03-JUN-2002; 2002US-0385691P.
XX PA (GENE-) GENE-CRAFT INC.
XX PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
XX DR WPI; 2003-801317/75.
XX PT New binding domain-immunoglobulin fusion protein, useful for treating a
XX PT subject having or suspected of having a malignant condition or a B-cell
XX PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX PS Disclosure; SEQ ID NO 176; 157pp; English.
XX CC The invention relates to a binding domain-immunoglobulin fusion protein
XX CC comprising a binding domain polypeptide that is fused to an
XX CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
XX CC CH2 constant region polypeptide that is fused to the hinge region
XX CC polypeptide, and an immunoglobulin heavy chain CH3 constant region
XX CC polypeptide that is fused to the CH2 constant region polypeptide. The
XX CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
XX CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
XX CC region polypeptide; derived from (a) having 3 or more cysteine residues;
XX CC where the mutated human IgG1 immunoglobulin hinge region polypeptide
XX CC contains 2 cysteine residues, where the first cysteine is not mutated; a
XX CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
XX CC (a) having 3 or more cysteine residues, where the mutated human IgG1
XX CC immunoglobulin hinge region polypeptide contains no more than one
XX CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
XX CC polypeptide, derived from (a) having 3 or more cysteine residues; where
XX CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains
XX CC no cysteine residues. The binding domain-immunoglobulin fusion protein is
XX CC capable of at least one immunological activity comprising antibody
XX CC binding domain polypeptide is capable of specifically binding to an
XX CC antigen. Also included are an isolated polynucleotide encoding the
XX CC binding domain-immunoglobulin fusion protein, a recombinant expression
XX CC construct comprising the polynucleotide (operably linked to a promoter),
XX CC a host cell transformed or transfected with a recombinant expression
XX CC construct, producing the binding domain-immunoglobulin fusion protein, a
```

CC pharmaceutical composition comprising the binding domain-immunoglobulin
 CC fusion protein or polynucleotide and a carrier, and treating a subject
 CC having or suspected of having a malignant condition or a B-cell disorder.
 CC The binding domain-immunoglobulin fusion protein is useful for treating a
 CC subject having or suspected of having a malignant condition or a B-cell
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
 CC sclerosis or autoimmune disease. The present sequence is a binding domain
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The
 CC sequence data for this patent formed part of the printed specification
 CC and is also available in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
 CC identified the sequences in the printed specification by their SEQ ID
 CC number therefore none of the sequences can be explicitly identified.
 CC
 XX
 SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 7; Length 1153;
 Best Local Similarity 99.2%; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
 Db 17 ENLDTENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMVLSGLSLAATTSPPQLLAGPTVHQTCSNTYVKGCLFLPGSNLRQOPQK 120
 Db 77 RLQVPVEAVNMVLSGLSLAATTSPPQLLAGPTVHQTCSNTYVKGCLFLPGSNLRQOPQK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEQLKSKTLFSLMQYSEEF 180
 Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKWVSTVMEQLKSKTLFSLMQYSEEF 196
 QY 181 RHFTPEKEQNNPNRSLKPIPTOLLGRHTATGLRKVRELPNIINGARQNAFKILFLL 240
 Db 197 RHFTPEKEQNNPNRSLKPIPTOLLGRHTATGLRKVRELPNIINGARQNAFKILVVI 256
 QY 241 TDGEKEGDPGLVEDVPELDREGVIRYVGVGDAPSEKSRQELNVTASKPRDHVFOIN 300
 Db 257 TDGEKEGDPGLVEDVPELDREGVIRYVGVGDAPSEKSRQELNVTASKPRDHVFOIN 316
 QY 301 NFEALKTIONQREKIFAIEGTQTGSSSFHEHMSQEGFSAATISNGPLLSVGSYDMAG 360
 Db 317 NFEALKTIONQREKIFAIEGTQTGSSSFHEHMSQEGFSAATISNGPLLSVGSYDMAG 376
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMPR 420
 Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMPR 436
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 480
 Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVEOTRGQSVCP 496
 QY 481 PRGQARWQCDVAVLGEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYL 540
 Db 497 PRGQARWQCDVAVLGEQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYL 556
 QY 541 HGTSGSGISPSHSORTAGSKLPRLOYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 Db 557 HGTSGSGISPSHSORTAGSKLPRLOYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
 QY 601 PVLVRKAIIMEFNPVAVRVNFCNDQVKGKAGEVRVCLHVQKSTRDLRSGQISQVVT 660
 Db 617 PVLVRKAIIMEFNPVAVRVNFCNDQVKGKAGEVRVCLHVQKSTRDLRSGQISQVVT 676
 QY 661 YDLALDGRPHGRAVNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
 Db 677 YDLALDGRPHGRAVNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
 QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLISITPSFMSLCLVVG 780
 Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLISITPSFMSLCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWRLACBSASSTEV 840
 Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSORSQSWRLACBSASSTEV 856
 QY 841 SGALKSTSCSINHIPIFPENSEVFNTFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
 Db 857 SGALKSTSCSINHIPIFPENSEVFNTFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
 QY 901 QLELPVKYAYVMVTVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISIVFLVPV 960
 Db 917 QLELPVKYAYVMVTVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISIVFLVPV 976
 QY 961 RLNQTVWDRPQVTFSENLSSTCHTERLPSSHSDFLAELRKAPVNVCSIAVCORIQCDDIP 1020
 Db 977 RLNQTVWDRPQVTFSENLSSTCHTERLPSSHSDFLAELRKAPVNVCSIAVCORIQCDDIP 1036
 QY 1021 FFGIOEFNATLKGNLKNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGOGAFVRSQTET 1080
 Db 1037 FFGIOEFNATLKGNLKNSLFDWYIKTSHNHLILVSTAEILFNDSVFTLLPGOGAFVRSQTET 1096
 QY 1081 KVEPFVFPNPLPIVGVSSVGGLLLLALITAALYKLGFFKRYKDMMSGGPPGAEPPQ 1137
 Db 1097 KVEPFVFPNPLPIVGVSSVGGLLLLALITAALYKLGFFKRYKDMMSGGPPGAEPPQ 1153
 RESULT 7
 AAR04136
 ID AAR04136 standard; protein; 1153 AA.
 AC AAR04136;
 XX
 XX 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 07-SEP-1990 (first entry)
 XX
 DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
 XX
 KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
 KW non-specific defence system; integrin gene superfamily.
 XX
 OS Synthetic.
 XX
 FH Key
 FT Location/Qualifiers
 FT 1..16
 FT /label= signal_peptide
 FT Modified-site
 FT 86..88
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 240..242
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 391..393
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 469..471
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 693..695
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 697..699
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 735..737
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 802..804
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 881..883
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 901..903
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 912..914
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 941..943
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 947..949
 FT /label= putative N-glycosylation site
 FT Modified-site
 FT 979..981
 FT /label= putative N-glycosylation site

FT Modified-site 994..996
 FT /label= putative N-glycosylation site
 FT Modified-site 1022..1024
 FT /label= putative N-glycosylation site
 FT Modified-site 1045..1047
 FT /label= putative N-glycosylation site
 FT Modified-site 1051..1053
 FT /label= putative N-glycosylation site
 FT Modified-site 1076..1078
 FT /label= putative N-glycosylation site
 FT Region 1106..1134
 FT /label= putative_transmembrane_region
 XX
 FN EP264690-A.
 XX
 XX 25-APR-1990.
 XX
 XX 17-AUG-1989; 89EP-00115159.
 XX
 XX 23-AUG-1988; 88US-00235353.
 PR 09-MAR-1989; 89US-00321239.
 XX
 XX (DAND) DANA FARBER CANCER INST INC.
 XX
 XX Springer TA, Corbi A;
 XX WPI; 1990-125938/17.
 DR N-PSDB; AAQ04043.
 XX
 XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
 PT inflammation and viral infections, and in diagnosis.
 XX
 XX Disclosure; Page ?; 3pp; English.
 XX
 XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
 CC -MAR-2003 to correct PA field.)
 CC
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key
 CC and pages
 XX
 XX Sequence 1153 AA;

Query Match 99.3%; Score 5842; DB 2; Length 1153;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 QY 1 FNLDTENAMTQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDSYSGCEPI 60
 DB 17 FNLDTENAMTQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDSYSGCEPI 76
 QY 61 RLQPVVEAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVVKGLCFGLFNSLRQPOK 120
 DB 77 RLQPVVEAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVVKGLCFGLFNSLRQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEVSTVMEQLKSKTFLSLMQYSEEP 180
 DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEVSTVMEQLKSKTFLSLMQYSEEP 196
 QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILL 240
 DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILLVVI 256
 QY 241 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAFSEKSRQELNTVASKPRDHVFQIN 300
 DB 257 TDGKFGDPLGVEDVPEADREGVIRYVIGVDAFSEKSRQELNTVASKPRDHVFQIN 316
 QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPILLSTVGSYDWAG 360
 DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPILLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRYQSLVLGAPRYOHIHGLVAMFR 420
 DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRYQSLVLGAPRYOHIHGLVAMFR 436
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 496
 QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTVAIGAPGEDNRGAYLTF 540
 DB 497 PRGQARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTVAIGAPGEDNRGAYLTF 556
 QY 541 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAGQHVLRLSRQ 600
 DB 557 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAGQHVLRLSRQ 616
 QY 601 PVLRVKAIMEFNPREVARNVFECDQVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 660
 DB 617 PVLRVKAIMEFNPREVARNVFECDQVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 676
 QY 661 YDLALDSGRPHSAVFNETKNSRRTQVTLGTCTETLKLQLPNCIEDPVSPIVLRNF 720
 DB 677 YDLALDSGRPHSAVFNETKNSRRTQVTLGTCTETLKLQLPNCIEDPVSPIVLRNF 736
 QY 721 SLVGTPLSAFNLRPVLAEDAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFNLRPVLAEDAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLQSVKYSTLQNRQSRQWRLACESASTEV 840
 DB 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLQSVKYSTLQNRQSRQWRLACESASTEV 856
 QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
 DB 857 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 916
 QY 901 QLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMQHQYQVSNLQSRSLPISLVLFPV 960
 DB 917 QLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMQHQYQVSNLQSRSLPISLVLFPV 976
 QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
 DB 977 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
 QY 1021 PFGIQEEFNATLKNLSFDMVYKTSNHLIIVSTABILFNDSVFTLLPGOGAFVRSQET 1080
 DB 1037 PFGIQEEFNATLKNLSFDMVYKTSNHLIIVSTABILFNDSVFTLLPGOGAFVRSQET 1096
 QY 1081 KVEPFEPVNPPLIVGSSVGLLALLITALYKLGFFKRYQKQKMMSEGGPPGAPQ 1137
 DB 1097 KVEPFEPVNPPLIVGSSVGLLALLITALYKLGFFKRYQKQKMMSEGGPPGAPQ 1153
 RESULT 8
 ADM99589
 ID ADM99589 standard; protein; 1152 AA.
 XX
 AC ADM99589;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Human integrin alphaM subunit precursor protein.
 XX
 KW integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
 KW osteopathic; cytostatic; immunosuppressive; anti-inflammatory;
 KW neuroprotective; antisticking; immunotherapy; inflammatory;
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
 KW alphaM.
 XX
 OS Homo sapiens.
 XX

Key Location/Qualifiers
Misc-difference 965
/note= "Encoded by CCC"

WO2004007530-A2.

22-JAN-2004.

17-JUL-2003; 2003WO-US022301.

17-JUL-2002; 2002US-0396783P.

17-JUL-2002; 2002US-0396790P.

11-SEP-2002; 2002US-0410135P.

(BLOO-) CENT BLOOD RES INC.

Springer TA, Takagi J;

WPI; 2004-122877/12.

N-PSDB; ADM99588.

Novel modified integrin protein having extracellular domains of integrin alpha and beta subunits or integrin alpha1 and beta3 subunit, useful for treating integrin mediated disorders.

Disclosure; SEQ ID NO 4; 232pp; English.

The invention relates to a novel isolated or recombinant modified integrin protein having extracellular domains of integrin alpha and beta subunits where one of the subunits has one or more mutations, an altered surface feature or an amino acid substitution or internal deletion, extracellular domains of the integrin beta subunit that comprise a mutation that alters a non-cysteine residue to cysteine or extracellular domains of integrin alpha and beta subunits. The polypeptide of the invention demonstrates antiproliferative, thrombolytic, anticoagulant, osteoprotective, cytostatic, immunosuppressive, antiinflammatory, neuroprotective and anticircling activities and may be useful for immunotherapy in order to prevent or treat an integrin-mediated disorder such as an inflammatory disorder, an autoimmune disorder, thrombosis, cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple sclerosis. The current sequence is that of the human integrin alpha subunit precursor protein of the invention.

Sequence 1152 AA;

Query Match 99.2%; Score 5836.5; DB 8; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
61 RLQVPVEANMGLSLAANTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPPQK 120
77 RLQVPVEANMGLSLAANTSPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQPPQK 136
121 FPEALRGCPQEDSIAFLDGSIIIPHPFRMKKEWVSTVMEQLKSKTFLSLMQVSEEF 180
137 FPEALRGCPQEDSIAFLDGSIIIPHPFRMKKEWVSTVMEQLKSKTFLSLMQVSEEF 196
181 RIHFTKEFQNNPNRSLKIPITQLGRTHATGLRKVVRELNTINGARKNAFKLFL 240
197 RIHFTKEFQNNPNRSLKIPITQLGRTHATGLRKVVRELNTINGARKNAFKLFL 256
241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVGDAFSEKSRQELNTVASKPRDHVFQIN 300
257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVGDAFSEKSRQELNTVASKPRDHVFQIN 316
301 NFEALKTIONQREKIFALEGTQTGSSSSFEHMSQEGFSAITNSGPLLLSTVGSYDMAG 360
317 NFEALKTIONQREKIFALEGTQTGSSSSFEHMSQEGFSAITNSGPLLLSTVGSYDMAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGGOVSCPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEOTRGGOVSCPL 496
QY 481 PRGORAPWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDAIGAAGBEDNRGAYVLF 540
DB 497 PRG-RARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDAIGAAGBEDNRGAYVLF 555
QY 541 HGTSGSGISPSHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 556 HGTSGSGISPSHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 615
QY 601 PVLVRKAIMEFNPREVARNVFECDVWVGKEAGEVRVCLHVOKSTRDRIRREGIOQSIVT 660
DB 616 PVLVRKAIMEFNPREVARNVFECDVWVGKEAGEVRVCLHVOKSTRDRIRREGIOQSIVT 675
QY 661 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720
DB 676 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 735
QY 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFPFKXKCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFGNLRPVLAEDAQRLFTALPFPFKXKCGNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGSDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 840
DB 796 GPREFNVTVVRNDGSDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 855
QY 841 SGALKSTCSINHIPIFPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 900
DB 856 SGALKSTCSINHIPIFPENSEVTFNITFDVDSKASLGNKLLKXANTVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVYVWVTSHGVSXKVLNFTASENTSRVWQHOYQVSNLQORSPLISLFLVPV 960
DB 916 QLELPVKYAVYVWVTSHGVSXKVLNFTASENTSRVWQHOYQVSNLQORSPLISLFLVPV 975
QY 961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSTAVCORIOCDIP 1020
DB 976 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSTAVCORIOCDIP 1035
QY 1021 FFGIOBEFNATLKGNSFDWYIKTSHNLLIIVSTABILENDSVFTLLPQGGAFVRSQTET 1080
DB 1036 FFGIOBEFNATLKGNSFDWYIKTSHNLLIIVSTABILENDSVFTLLPQGGAFVRSQTET 1095
QY 1081 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
DB 1096 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1152

RESULT 9
ADP12435
ID ADP12435 standard; protein; 1152 AA.
XX
AC ADP12435;
XX
DT 12-AUG-2004 (first entry)
XX
DE Protein encoded by mRNA of the invention #45.
XX
KW transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX
OS Homo sapiens.
XX
PN WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
PF 24-APR-2003; 2003WO-US012946.

XX PR 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
PI WPI; 2004-400724/37.
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
XX rejection, in an individual, comprises detecting the expression level of
XX the genes.
XX Claim 65; SEQ ID NO 2444; 1762pp; English.
XX The present invention relates to diagnosing or monitoring transplant
XX rejection, e.g. cardiac or kidney transplant rejection, in an individual
XX comprises detecting the expression level of one or more genes. The
XX methods, system and kits are useful in diagnosing or monitoring
XX transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
XX islet, lung, bone marrow or stem cell transplant rejection,
XX xenotransplant rejection or mechanical organ replacement rejection, in an
XX individual. The method is also useful in assessing the immune status of
XX an individual. The methods are also useful in diagnosing and monitoring
XX diseases that involve the immune system, e.g. rheumatoid arthritis,
XX lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
XX viral, bacterial or fungal infection. The present sequence represents a
XX protein that is encoded by the mRNA of the invention.
XX SQ Sequence 1152 AA;
Query Match 99.1%; Score 5829.5; DB 8; Length 1152;
Best Local Similarity 99.0%; Pred. NO. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;
QY 1 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDSYSGCEPI 60
DB 17 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDSYSGCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVYKGLCFGLFSGNLROQPOK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVYKGLCFGLFSGNLROQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIIPHDFFRMKEWVSTVMQELKSKXTLFLSMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGIIIPHDFFRMKEFVSTVMEQLKSKXTLFLSMQYSEEF 196
QY 181 RIHTFFKFPQNNPNRSLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILFL 240
DB 197 RIHTFFKFPQNNPNRSLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVIGVDAPRSKQELNTVASKPDRHVFQIN 300
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSKQELNTVASKPDRHVFQIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKSTKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKSTKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGWNESNANVKGTOIGAYFASLCSVDVDNGSGTDLVLI GAPHYEQTGRGQVSVCP 480
DB 437 QNTGWNESNANVKGTOIGAYFASLCSVDVDNGSGTDLVLI GAPHYEQTGRGQVSVCP 496
QY 481 PRGQBARWCDAVLVYGEQQQWGRFGAALTVLGDVNGDKLTDAVIGAPGEEDNRGAVYLF 540
DB 497 PRG-RARWCDAVLVYGEQQQWGRFGAALTVLGDVNGDKLTDAVIGAPGEEDNRGAVYLF 555

QY 541 HGTSGGISPSHSORIASGKLSPLQYFGOSLSGGQDLTMDGLVLDLTVGAGQHVLRLRSQ 600
DB 556 HGTSGGISPSHSORIASGKLSPLQYFGOSLSGGQDLTMDGLVLDLTVGAGQHVLRLRSQ 615
QY 601 PVLVRKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLRREGIOQSVVT 660
DB 616 PVLVRKAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLRREGIOQSVVT 675
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720
DB 676 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 735
QY 721 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPFEXKNCNDNICQDDLSITFFSMDLCLVVG 780
DB 736 SLVGTPLSAFNGNRPVLAEDAQRLLFTALFPFEXKNCNDNICQDDLSITFFSMDLCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLONQSRQSRWRLACESASSTEV 840
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLONQSRQSRWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960
DB 916 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 975
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035
QY 1021 FFGIOEEFNATLKNLSFDWIKTSHNHLITVSTAEILFNDVSFTLLPGQGAFFVRSQTET 1080
DB 1036 FFGIOEEFNATLKNLSFDWIKTSHNHLITVSTAEILFNDVSFTLLPGQGAFFVRSQTET 1095
QY 1081 KVEPFEPVNPPLPLTVGSSVGGLLLELALITAAALYKLGFFKRYKQKMMSEGGPPGAEPPQ 1137
DB 1096 KVEPFEPVNPPLPLTVGSSVGGLLLELALITAAALYKLGFFKRYKQKMMSEGGPPGAEPPQ 1152
RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX AC ADP44061;
XX DT 09-SEP-2004 (first entry)
XX DE Human CD11C protein SEQ ID NO:14.
XX KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
XX anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX OS Homo sapiens.
XX FN WO2004053094-A2.
XX PD 24-JUN-2004.
XX PF 08-DEC-2003; 2003WO-US039208.
XX PR 06-DEC-2002; 2002US-0431522P.
XX PA (PPDP-) PPD DEV LP.
XX PI Dunn SJ, Holzmayer TA;
XX DR WPI; 2004-480928/45.
XX DR N-PSDB; ADP44060.
XX PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

a human host cell useful for preventing and/or treating HIV infection, by identifying an inhibitor of a cell surface polypeptide such as CXCR-4.
 Claim 1; SEQ ID NO 14; 133pp; English.

The present invention describes a method (M1) for identifying an inhibitor of HIV entry into a human host cell. (M1) comprises identifying an inhibitor of a cell surface polypeptide selected from CXCR-4 (352 amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10, ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11c (1163 amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69 (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077), CD11c (1299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids, SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36, ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide as described above in a human host cell preventing HIV entry into the human host cell; (2) a pharmaceutical composition comprising an inhibitor of (1) and a carrier; and (3) conferring resistance to HIV infection in an individual, by administering the pharmaceutical composition of (2). An inhibitor of HIV entry has anti-HIV and virucide activities, and can be used as an HIV uptake inhibitor. (M1) is useful for identifying protective compounds that inhibit entry of HIV into cells, useful for the prevention and/or treatment of HIV infection. The present sequence represents human integrin alpha X (CD11c), which is used in the exemplification of the present invention.

Seq Sequence 1163 AA;

Query Match 59.0%; Score 3473; DB 8; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 1.4e-281;
 Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

1 FNLDTENAMTFQENARGFQSVQLOGSVVVGAPQEIIVANRGSLYQCDYSTGCEPI 60
 20 FNLDTTELTAFRVDSDAGFQSVVQVYANSVVVVGAPQKITAANQTGGLYQGYSTGCEPI 79
 61 RLQVPEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVYKGLCFLLFGSNLRQOPQK 120
 80 GLQVPEAVNMSLGLSLASTSPQLLACGPTVHCECRNXYLTGLCFLLGPT-QLTQR 137
 121 FPEALRGCPQEDSDIAFLIDSGSIIPHPFRKKEWSTVMEQLKSKTILFSLMQVSEEF 180
 138 LPVSRQCEPQEQDIFVLIDSGSISRNFAWNVFRAVISQFQRPSTQFSLMQFSNKF 197
 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKRVVRELFNITNGARKNAFKILFL 240
 198 QTHFTPEFRSSNPLSLASVHQLQGFTYTATAIQNVVHRLPHASVGARRDAKILIVI 257
 241 TDGEKEDPLGYEDVPELDREGVIRVIGVDAFSEKSRQELNVTASKPPRDHVFQIN 300
 258 TDGKKGSDLDYKDVFPADMAAGIIRIYAGVLAQFQNRNSKELNDIASKPSQEHIFKVE 317
 301 NFPAKTIQNLQREKIFALEGTGTGSSSEPEHMSQEGFSAATNSGPLLSTVGSVDWAG 360
 318 DFDALXIQNLQREKIFALEGTGTGSSSEPEHMSQEGFSAATNSGPLLSTVGSVDWAG 377
 361 GVFLYTSKESKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVGLAPRYQHILGVAMFR 420
 378 GAFLYPPNNSPTFINNSQENVMDRDSVLYGSTELALWKGVQSLVGLAPRYQHITGKAVIT 437
 421 QNTGMHESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
 438 QVSRQWRMAEVTGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 497
 481 PRQARWOCDAVLYGEQGPWCFGAALTVLGDVNGDKLTQVYICAPGEEDNRGAVLYF 540
 498 PRGWR-RWNCDAVLYGEQGPWCFGAALTVLGDVNGDKLTQVYICAPGEEDNRGAVLYF 556
 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDITVGAQGHVLLRQ 600

Db 557 HGVLGPSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDITVGAQGHVLLR 616
 QY 601 PVLRVKALMEFNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
 Db 617 PVLWVGSMQIPAEIPRSAFECEHQVQVSEGLVQSNICLYDKRKNLLGSRDQSSVT 676
 QY 661 YDLALDSGRPHSRVAFVETKNTSRQVIGLGTCTETLKLQLPNCIEDPVSPIVIRLNF 720
 Db 677 LDALDLCRLSPRATFOETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVTPIRLNF 736
 QY 721 SLVGTPLSAGFNRLPVLAEQAORLFTALFPFCKNCNDNICODDLISITFFSMDSCLVVG 780
 Db 737 TLVGPPLAFNRDLRPLMLAADAQRYFTASLPFCKNCADHICQDNLGHSFSPFGLKSLIV 796
 QY 781 GPRFNVTVTVNRDGEDSYRTQVTFPPFLDLVYRKVSTLQNRQSRQWRRLACESASSTEV 840
 Db 797 SNLELNAEVMVNDGEDSYGTTITFHPAGLSYRYVAEGQKQGLRSLHLTCDAPVG-- 854
 QY 841 SGALKSTSCINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 Db 855 SQGTWSTSCRINHLIFRGAQITFLATFDVSPRAVLGDRLLLTANVSENNTPTSKTTF 914
 QY 901 QLELPVYAVVYVTVSHGVSTKYLNFAS-ENTSRVMOHQYVSNLQORSLPISLVLPV 959
 Db 915 QLELPVYAVVYVTVSHGVSTKYLNFSESEKSHVAMHRYQVNNLQORLDPVSNFVWF 974
 QY 960 VRLNQTVINDRPQVTFSENLSTCHTKERLPFSDHFLAELRKAPVNVCSIAVCORICDI 1019
 Db 975 VELNQEAVMVDVESHVHPQNPDLRCSSEKIAPPASDFLAHQKPFVLDCCSIAGCLRFCDV 1034
 QY 1020 PFGIOBEFNATLKNLSFDWYIKTSHNHLIYVSTABILEFNDVSFTLLPQGGAFVRSQTE 1079
 Db 1035 PSESVOBELDFTLKNLSFGWVRQILQKVVSVVAEITFDTSVYSQLPQGEAFMRAQT 1094
 QY 1080 TKVEPEVNPPLIYGVSSVGGLLILALITAAALYKLGFFKQYKDMSE 1128
 Db 1095 TVLEKYKVNPTPLIYGVSSVGGLLILALITAVLYKGVGFKRYKEMMEE 1143

RESULT 11
 ADN02004
 ID ADN02004 standard; protein; 1163 AA.
 AC ADN02004;
 XX ADN02004;
 DT 17-JUN-2004 (first entry)
 XX Human inflammatory bowel disease marker - CD11c protein.
 DE marker gene; inflammatory bowel disease; IBD; PcgammaR IIIa;
 XX PcgammaR IIib; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);
 KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;
 KW Crohn's disease.
 XX Homo sapiens.
 OS JP2004065120-A.
 PN 04-MAR-2004.
 XX 07-AUG-2002; 2002JP-00229705.
 PF 07-AUG-2002; 2002JP-00229705.
 PR 07-AUG-2002; 2002JP-00229705.
 XX (SUMU) SUMITOMO SEIYAKU KK.
 XX WPI; 2004-209124/20.
 DR N-PSDB; ADN01990.
 DR Novel marker gene of inflammatory bowel disease (IBD) comprising base
 PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIib, Mig, NRG-2,
 PT hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic

PT agent for IBD.
XX Example; SEQ ID NO 27; 151pp; Japanese.
PS
XX
CC The invention comprises marker genes for inflammatory bowel disease (IBD)
CC - FcgammaR IIa, FcgammaR IIb, Mig, NRG-2, hexokinase, HM74, REG III,
CC LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC The DNA and protein sequences of the invention are useful in the
CC detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC disease). The present amino acid sequence represents an IBD marker
CC protein of the invention.
XX
XX
XX Sequence 1163 AA;
Query Match 58.8%; Score 3459; DB 8; Length 1163;
Best Local Similarity 60.9%; Pred. No. 2.1e-280;
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;
QY 1 FNLDENAMTFQENARGGQGVVQVQGSRRVVGAPQEIIVAAANQSGSLYQCDSYSGCEPI 60
Db 20 FNLDTEELTAFRVDSAGFSDSVVQVANSVVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVFEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQQPQK 120
Db 80 GLQVPEAVNMSLGLSLAATTSPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 PPEALRGCPQEDSDIAFLIDSGSII PHDPRMKEWSTVMEQLKSKTLESLMOYSEEF 180
Db 138 LFSVRQECPRQEQDIFVLIDSGSISRRNFATMNFVRAVISQFQPSQSLMWFQSNKF 197
QY 181 RIHFFTEKFNPNPRSLIKPITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240
Db 198 QHTHFFERTNSPLSLASVHQLOGFTYATAIQNVVHRLFHASYGARRDATKILIVI 257
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVQIN 300
Db 258 TDGKKEGSLDYKQVIMADAAAGIIRYAIGVLAFONRNWKNELNDIAKPSQEHIFKVE 317
QY 301 NFEALKTTQNLQREKIFAIEGTQTSSTSSPHEMSQEGFSAATISNGPLLSVGSYDAG 360
Db 318 DFDALDKIQNLQREKIFAIEGTQTSSTSSFELEMAQEGFSAFTPDGVLGAVGFTWSG 377
QY 361 GVELTSKESKPFINNTVDSMDNAYLGAIAAIIILNRKVSILVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSGVNDNRDSYLGYSTELALWKGVQSLVGLGAPRYQHTGAVIFT 437
QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCLP 480
Db 438 QVSRQWRMKAETGTOIGSYFCASLCSVDVDTGSDTDLVLI GAPHYEQTRGGQVSVCLP 497
QY 481 PRQVRRQCDVLYGEOQWGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAVILF 540
Db 498 PRGWR-RWMCDAVLYGEOQWGRFGAALTVLGDVNGDKLTDVIGAPCEENRGAVILF 556
QY 541 HTSGSGISPSHSQRIAGSKLPRLOYFCQSLSGGQDLTMDGLVDLTGCAQGHVLLLRSQ 600
Db 557 HGVLGSPISPSHSQRIAGSKLPRLOYFCQALSGGQDLTQDGLVDLAVGARGQVLLLR 616
QY 601 PVLRYKAIMFNPREVARNVFCDQVVKKEBAGEVRVCHLVQKSTRDRRLRGQOSVVT 660
Db 617 PVLVGVSMQFIPAEIPRSAFECEQVSEQTLVQSNICLYIDKRSKNLGRDLQSSVT 676
QY 661 YDLALDGRPHARVFNENKSTRTOVLGHTQTCETLKLQLPNCIEPVPFIVLRNF 720
Db 677 LDALDPGLRSPRAIFQETKNSLSRVRLGLKAHCENFNLLPSCVDSVPTILRLNF 736
QY 721 SLVGTPLSAFGLNRLVLAEDAQRLFTALFPFKEKNGNDNICQDLSITFSFMSLCLVVG 780
Db 737 TLVGKPLLAFLAFLNRLPMLAALQRYPTASLPFEKNCQADHICQDNLGISFSPGLKSLVG 796
QY 781 GPREFNVTVRNDGSDSTRTQVTFPFLDLSVRKYSTLTQNRQSRWSLACESASSTEV 840
Db 797 SNLELNAEVMVNDGSDSYGTTITTFSHPAGLSYRYVAEGKQKQQLRSLHLTCDSPVGV-- 854

841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPTKTEF 900
855 SQGTWSTSCSINHPIFRGGAQITFLATFDVSPRAVLGDRLLLTANVSSENNTPTSKTTF 914
901 QLELPVKYAVYVMVTSHTKYLINFTAS-ENTSRVMOHQVQVSNLCQSLPISLVFLVP 959
915 QLELPVKYAVYVTVSSHEQTKYLNFSSEBESHVAMHRYQVNNLQORLPVSNFWVP 974
960 VRLNQTVIDRPOVTFSENLSSCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIODCI 1019
975 VELNQEAVMVDVEVSHQPNFSLRCSSEKIAPPASDFLAHQKNPVLDCSTAGCLFRCDV 1034
1020 PFFGIOEFNATLKNLSFDWYIKTSHNHLIIYSTAELFNDSVFTLLPQOGAFVRQOTE 1079
1035 PFSVQBELDTLLKGNLSFGWVRQILQKVSVVSVABITDTSVYSLPGQEAFFMQATT 1094
1080 TKVPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMWSE 1128
1095 TVLEKYKVNPTPLIVGSSIGGLLLALITAVLYKVGFFKRYKEMMEE 1143

RESULT 12
ADQ17510
ID ADQ17510 standard; protein; 1163 AA.
XX AC ADQ17510;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
XX of a gene in a first soft tissue sample and a normal soft tissue sample
XX and comparing the gene expression, also useful in treating soft tissue
XX sarcoma.
XX
XX Example 2; SEQ ID NO 327; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
XX which comprises obtaining a first soft tissue sample from an individual
XX and a normal soft tissue sample from the same or different individual,
XX determining the expression of a gene in both samples and comparing the
XX expression of the gene in both soft tissue samples, where a higher level
XX of protein expression in the first soft tissue sample indicates the
XX presence of soft tissue sarcoma. The method of the invention has
XX cytostatic applications and may be useful for detecting soft tissue
XX sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
XX protein of the invention. The current sequence is not shown within the
XX specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 1163 AA;
SQ

CC Clone lambda X47 was isolated from a cDNA library constructed from total
 CC RNA extracted from phorbol myristate acetate stimulated HL-60
 CC myelomonocytic cells. The library was screened with oligonucleotide
 CC probes based on tryptic peptide fragments of p150.95. The sequence can be
 CC attached to appropriate control elements and expressed in prokaryotic and
 CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
 CC infection because it interacts with ICAM-1 and inhibits cell-virus
 CC attachment. It can also be used as an anti-inflammatory agent. See also
 CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
 CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 1163 AA;

Query Match 58.6%; Score 3450; DB 2; Length 1163;
 Best Local Similarity 60.7%; Pred. No. 1.2e-279;
 Matches 685; Conservative 142; Mismatches 296; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARFGVGSGVQLOGSRVVGAPQIVAAQNRGSLYQCDYSGSCPEI 60
 Db 20 FNLDTBELTAFRVDSAGFGSVVQYANSWVVGAPQKIATANQTGGLYQCGYSTGACEPI 79

QY 61 RLQVPVBAVNSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFGLFSNLRQDPQX 120
 Db 80 GLQVPPBAVNSLGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCFLLGPT--QLTQR 137

QY 121 FPEALRGCPQEDSIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180
 Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMNFRVAVISQFQSTQFSLNQFSNKF 197

QY 181 RIHETVEFQNNPNSRIKIPITOLGRTHTATGLRKVVRLEPNTGARKNAKFIKILL 240
 Db 198 QTHETFEFRFTSNPLSLASVHQGGTYTATAIQNVHRLPHASVGARDAYKLLIVI 257

QY 241 TDGEKFGDPLGYDVIPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPRDRHVFQIN 300
 Db 258 TDGKKEGDSLDYKDVIPMADAAGIIRYVIGVGLAFQNRNSWKELNDIASKPSQEHIFKVE 317

QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSPHEMSQEGFSAITNSGELLSTVSGYSOWAG 360
 Db 318 DFDALKDQNLQREKIFAIEGTQTGSSSPHEMSQEGFSAITNSGELLSTVSGYSOWAG 377

QY 361 GVELYTSKEKSTFNTMTVDSDMDAYLVGAAAIILNRNQSLSVLGAPRYQHIGLVAMFR 420
 Db 378 GAFIYPPNMSPTFFINNSQVMDRDSYLGYSYELALWKGVSQSLVGLGAPRYOHTGKAVIFT 437

QY 421 QNTGMWESNANVKGTCIGAYFCASLCSVDVDSNGSTDVLVLCAPHYVEQTRGGQVSVCP 480
 Db 438 QVSRQWRMKAETVGTQIGSYFCASLCSVDVDSNGSTDVLVLCAPHYVEQTRGGQVSVCP 497

QY 481 PRGQARQCDAVLYGEOGPWGRFGAALTVLGVNGDKLTDVAIGAPGSDNRGAVYLF 540
 Db 498 PRGWR-RWMCDAVLYGEOGHPWGRFGAALTVLGVNGDKLTDVWVIGAPGEBENRGAVYLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVLTGVAQCHVLLRSQ 600
 Db 557 HGVLGPSISPSHSQRIAGSKLSSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR 616

QY 601 PVLRYKAJMEFNPREVARNVFECNDQVYKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
 Db 617 PVLWGVSMQFTFAPISAPFCREQVWSEQTLVQSNICLYIDKSKNLLGASRDLOQSSVT 676

QY 661 YDLALDGRPHSRVAFNETKSTRQTVLGHQTQCEILKLPNCIEDPPYSPVLRNLNF 720
 Db 677 LDALDGRPLSPRATFQETKNSLSRVRVLGKAHCENFNLLPSCVEDSVPTITRLNF 736

QY 721 SLVGTPLGAFGNLRPVLAEDAQRLFTALPFEKNCNDNICODDLSITFSFMSLCLVVG 780
 Db 737 TLVHGPLLAFLNLRPMLAALQRYFTASLPFEKNCADHIQDNLGISFSPFLKSLLVG 796

QY 781 GPREFNVTVTVNRDGEDSVRTQVTFPFLDLRYKRVKSTLQNRQSRWSLACESASSTEV 840
 Db 797 SNLELNAEVMWNWNGEDSVGTITTFSHPAGLSRYVVAEGQKQGLRSLHLTCDSA--PVA 854

QY 841 SGALKKTSINPIHPENSEVFNITFDVDSKASLGNKLLKANKYVTSNNMPTNKTEF 900
 Db 855 SQGTWTSCEINHLIRFGGAQITFLATFDVSPKAVLGDRLLLTANVSSENTERTSKTF 914

QY 901 QLELPVKYAVYVWVTVSHGVSTKYLVNTAS-ENTSRVWQHOYQVSNLQORSLSPLSLVLP 959
 Db 915 QLELPVKYAVYVTVSSHQFTKYLNFSESEKESHVAMHRYQVNNLQGRDLPSVSNPWP 974

QY 960 VRLNQTVWDRPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVNCIAVCQRIQCDI 1019
 Db 975 VELNQAVWMDVEVSHPNPNSLRCSSOKIAPPASDFLAHQKPNVLDGSLAGCLRFCDV 1034

QY 1020 PFTGIOBEFNATLKGMLSFWDYIKTSHNHLIIYSTAELFNDSVFTLLPQGGAFVRQOTE 1079
 Db 1035 PFSVQOEELDTLKGMLSGWVQILQKKVSVVSVAEITFDTSVYSQLPQGERFMAQT 1094

QY 1080 TKVEPFEVNPPLDILVGVSSVGGLLLLALITTAALYKLGFFKRYQKDWSE 1128
 Db 1095 TVLEKYKVNHPTELVGVSSIGALLLLALITAVLYKVGFFKRYQKEMME 1143

RESULT 14
 AAW65091
 ID AAW65091 standard; protein; 1163 AA.
 AC AAW65091;
 XX 28-SEP-1998 (first entry)
 DT Human Beta-integrin CD11c subunit protein.
 XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
 KW rheumatoid arthritis.
 OS Homo sapiens.
 XX US5728533-A.
 PN 17-MAR-1998.
 PD 07-JUN-1995; 95US-00485618.
 PF 23-DEC-1993; 93US-00173497.
 PR 05-AUG-1994; 94US-00286889.
 PR 21-DEC-1994; 94US-00362652.
 XX (ICOS-) ICOS CORP.
 XX Van Der Vieren M, Gallatin WM;
 XX WPI; 1998-206565/18.
 DR Screening assay for modulators of integrin binding - using immobilised or
 PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
 PS Example 5; Fig 1A-D; 106pp; English.
 XX This sequence represents a human beta-integrin CD11c subunit which is
 CC used to describe a method for identifying compounds that modulate the
 CC interaction of the beta-integrin alpha-d subunit with a binding partner
 CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
 CC d binding partner, one of which is immobilised and the other of which is
 CC labelled, in the presence of a test compound, and determining if the
 CC compound affects binding between the alpha-d polypeptide and alpha-d
 CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
 CC comprising the cytoplasmic, transmembrane or extracellular domain of
 CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
 CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
 CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
 CC and rheumatoid arthritis
 XX

SQ Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 2: Length 1163.

Best Local Similarity 60.6%; Pred. No. 1.8e-278;

Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY	1	FNLDTENAMTFQENARGFQSQSVVQLOGSRVVVGAPQBEIVAAANQRSGSLYQCDYGTSGCEPI	60
Db	20	FNLDTBELTAFRVDZSAGFGDSVVQVANSWVVVGAPQKIIAANQIGGLYQCXYGTGACPEI	79
QY	61	RLQVPVEAVNMSLGLSLAATSPCOLLACGPTVHQTCSENTYVYKGLCFLFGSNLRQPOK	120
Db	80	GLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHHEGRNMNLTGLCFLLGPT--QUTQR	137
QY	121	FPFALRGCPQEDSIAFLIDGSGSIIIPHDFRMKEWSTVMEQLKKSKTLFSLMQVSEEF	180
Db	138	LPVSRQCEPRQBDIVFLIDGSGSISRNFAFMNFVRAVISQPORSTQPSLMQPSNKE	197
QY	181	RTHFTFKEPQNNPRSLIKPITQLLGRTHTATGLRKVKVRELFININGARKNAFKLFL	240
Db	198	QHTFTEFBRFTSNPLSLASVHQLOQFTYTATAIQNVVHRLFHASYGARRDAIKLIVI	257
QY	241	TDGEXFGDPLGYEDVIPELDREGVIRYVIGVDFAFRSEKSRQELNTVASKPPDPHVQJIN	300
Db	258	TDGKKEGDSLDYKOVIPWADAAGIIRVAGVLAFQNRNSWKELINDTASQSEHIFKVE	317
QY	301	NFEALKTTQNLQRLREKIFAIEGTQTCSSSSFEHEMSQEGFSAAITSNGLTSTVGSWDAG	360
Db	318	DFDALKIQNLQKKEIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGPVLGAVGFSFWSG	377
QY	361	GVFLYTSKEKSTFINMTVRDSDMDDAYLGVAAAIIILNRVQSVLGNAPRYOHLGLVAMFR	420
Db	378	GAFLYPDMSPPTFINMSQENVMDRSLVGYSTELALWKGVQSVLVLGAPRYOHLGKAVIEI	437
QY	421	QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQGVSVCLP	480
Db	438	QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYYEOTRGQGVSVCLP	497
QY	481	PRGQARQWQDAVLGEGCQWPGRFGAALTVLGDNVNGDKLTDVAIGAPGEENRGAVYLF	540
Db	498	PRGMFR--RWMCDAVLGEGCQWPGRFGAALTVLGDNVNGDKLTDVVI GAPGEENRGAVYLF	556
QY	541	HGTSGSGISPSHSORIKAGSKLSPRLQYFGOSLGGQDLTMDGLDVTGCAQGHVILLRQ	600
Db	557	HGVLGSPISPSHSQRIAGSQLSRLQYFGQALSGQDLTDGLVDLAVARGQVILLRTR	616
QY	601	PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRLEGQIQSVVT	660
Db	617	PVLWVGVMQIPAEIPRSAFECEQVVSQTLVQSNICLIYIDKRSKNLLGSRDLQSSVT	676
QY	661	YDLALDGRPHSRVAVENETKNSRTOQVGLGTQTCETIKLQLPNCIEDPVSPIVLRNF	720
Db	677	LDLALAPGLRSPRAIFQETKNRSLSRVRVLGKAHCENFLLPSCVEDSVPTILRNF	736
QY	721	SLVGTPLSAFNLRPVLAEDAQRFLTALFPFEKNCNDNICDDLSITETFSNLSCLVVG	780
Db	737	TLVGKPELLAFNRLPMLAALAAQRVFTASLFPFNKCGADHICQDNLCISFSFPLKSLVVG	796
QY	781	GPREFNVTVVRNDGEDSYXTQVTFPPFLDLSYRKVSTLIQNRQSRQSMWLACESSTEV	840
Db	797	SNLELNABVWMDNGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHUTC--CSAPVG	854
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLIKANYTSENMPRINTKTEF	900
Db	855	SQGTWSTISCRINHIIIFRGGAQITFLATFDVSPKAVGLDRLLLTANVSSNNIPRTSKTIP	914
QY	901	QLELPVKYAVYVMTSHGVSTKYLNFITAS--ENTSRVMQHYQVSNLQGRSLPISLVFLVP	959
Db	915	QLELPVKYAVYIVVSSHEQTKYLNLFSESEKESHVAMHRYQVNNLQGRDLPVSNFVWP	974
QY	960	VLNQTVIWDPRPQVTFSENISSTCHTKERLPSHDSFLAELRKAPVNVCSIAVCORIQCDI	1019
Db	975	VELMQEAVMDVESHVQPNFSLRCSSEKTAAPPASDEFIAHIQKNPVLDSIAGCLFRCDV	1034

Matches		684;	Conservative	148;	Mismatches	291;	Indels	6;	Gaps	4;
Qy	1	FNLDTENAMTFOBNARFGSVVOLQGSRRVVVGAPOIIVAAORGSLYQCDYTGSCPEI	60							
Db	20	FNLDTBELTAFRVDASGDSVVQYANSWVVGAPQKIIAANOIGGILYQCGYSTGACEPI	79							
Qy	61	RLQVPVAVNMSLGLSAAATSPOLLACGPTVHQTCSNTYKVLGCLFGLGSLNLQOPOK	120							
Db	80	GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCLLGP--QUTQR	137							
Qy	121	FPBALRCPOBDSIAFLIDGSGSIIIPHDFRMEKWEVSTVMEQLKKSKTILFSLMQYSEEF	180							
Db	138	LPVSRQCPQEQDIFVLIDGSGSISRNATWNVFRAVISQORPSTQFSLMQFSNKF	197							
Qy	181	RIHFTFKPQNNPRSLIKPIITQLGRTHATGLRKVVRRELFNITNGARKNAPKILFLL	240							
Db	198	QTHFTFEFRRTSNPLSILASVHQLQGTFTATAIQNVVHRLFHASYGARRDAIKILIVI	257							
Qy	241	TDGEKFGDPIGYEDVIPELDREGVIRVVGVDARFSEKSRQELNTVASKPPRDHVFQIN	300							
Db	258	TDGKKEGSDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIPAVE	317							
Qy	301	NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPELLSTVGSYDWAG	360							
Db	318	DFDALKDIOQLKEKIFAIEGTETISSSSFELEWAQEGFSAVFTPDGVLGAVGSFTWSG	377							
Qy	361	GVFLYTSKEKSTFINMTRVSDMDNDAYLGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR	420							
Db	378	GAFLYPPNMSPTFINMSQENVDMRDSYLGSTELALMKGVQSLVILGAPRYOHIGKAVIFI	437							
Qy	421	QNTGWESNANVKTQICAYFGASLCSDVDNSGSTDILVLIGAPHYEQTRGGQVSVCPPL	480							
Db	438	QVSRQWRKAEVIGTQIGSYFGASLCSVDVDTGSTDILVLIGAPHYEQTRGGQVSVCPPL	497							
Qy	481	PRGORARQCDAVLVYGEQCPWGRFGAALTVLGDVNGDKLTDVAIGAAGEEDNRGAVYLF	540							
Db	498	PRGWR-RWNCDAVLVYGEQHPWGRFGAALTVLGDVNGDKLTDVWIGAPGEBEENRGAVYLF	556							
Qy	541	HGTSGSGTSPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ	600							
Db	557	HGVLGPSISPSHSQRIAGSQLSSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR	616							
Qy	601	PVLRYKATMENPREVAENVECNQDVVKGEAGVVRVCLHVOKSTRDLREGQIQSVVT	660							
Db	617	PVLWVGVMQFIPAEIPRSAFECEQVVBHQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676							
Qy	661	YDLALDGRPHSRVAFNETKNSRTRQTOVLGTQTCETLKLQLPNCIEDPVSPVILRLNF	720							
Db	677	LDLALAPGLSPRAIFQETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRLNF	736							
Qy	721	SLVGTPLSAPGNLRPVLAEDAQRLLTALPPEKNGCNDNIQDDLSITFSFMSLDCIYVG	780							
Db	737	TLVGKPLLAFLRNLRLPMLAALAQRVFTASLPEFKNGCADHIQDNIGISFSPGLKSLLVG	796							
Qy	781	GPREFNVTVVRNDCGDSVTRQVTFPPFLDLSYKRVSTLQORSORSWRLACESASSTEV	840							
Db	797	SNLELNAEVMWNDCGDSVTGTTITSHPAGLSYRVVAGQKQQLRSLHLTC--CSAPVG	854							
Qy	841	SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900							
Db	855	SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLORLILLIANVSSENNIPRTSKTIF	914							
Qy	901	QLELPVKAVYVTVTSHGVSTKYLNETAS-ENTSRVMQHOYQVSNLQORSPLISLVELVP	959							
Db	915	QLELPVKAVYTVWSHQFTKYLNFSESEKESHVAMHRYQVNNLQGRDLPLVSNFWVP	974							
Qy	960	VELNQTIVDRPOVTFSENLSTCTKRLPSHSPFLAELRKAPVNVNCIAVCQRIQCDI	1019							
Db	975	VELNQEAVMVDVEVHPQNPRLCSSEKIAFPASDFLAHIQKNPVLDCSIAGCLFRCDV	1034							
Qy	1020	PPFGIQEENFATLKNLSFDWIKTSHNHLIVSTAEIILFNDVSFTLLPGQAFVRSQTE	1079							
Db	1035	PSFSVQEELDFTLKNLSFGWVRQILQKKVSVVVAEIIIFDTSVYSQLPQEAFAFRAQTI	1094							

Qy 1080 TKVEPPEVNPPLPLIVGSSVGGILLALLALITAAALYKLGFFKQYKDMKSE 1128
Db 1095 TVLEKYVHNPIPLIVGSSIGGLLLALLITAVLYKVGFKQYKEMWEE 1143

Search completed: November 9, 2004, 12:08:19
Job time : 108.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds

(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAREPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5884	100.0	1137	10	US-09-902-481A-4
2	5857	99.5	1137	10	US-09-902-481A-3
3	5852	99.5	1153	9	US-09-350-259-3
4	5852	99.5	1153	10	US-09-902-481A-1
5	5852	99.5	1153	10	US-09-891-943-3
6	5852	99.5	1153	14	US-10-144-259-30
7	5852	99.5	1153	14	US-10-207-655-176
8	5851	99.4	1137	10	US-09-902-481A-5
9	5845	99.3	1137	10	US-09-902-481A-6
10	5836.5	99.2	1152	9	US-09-945-265-4
11	3459	58.8	1163	14	US-10-116-275-204
12	3436	58.4	1163	9	US-09-350-259-4
13	3436	58.4	1163	10	US-09-891-943-4

14	3401	57.8	1161	9	US-09-350-259-2
15	3401	57.8	1161	10	US-09-891-943-2
16	3385.5	57.5	1161	9	US-09-350-259-99
17	3385.5	57.5	1161	10	US-09-891-943-99
18	3224.5	54.8	1161	9	US-09-350-259-55
19	3224.5	54.8	1161	10	US-09-891-943-55
20	3212	54.6	1161	9	US-09-350-259-53
21	3212	54.6	1161	10	US-09-891-943-53
22	3203.5	54.4	1151	9	US-09-350-259-37
23	3203.5	54.4	1151	10	US-09-891-943-37
24	3191	54.2	1155	9	US-09-350-259-46
25	3191	54.2	1155	10	US-09-891-943-46
26	1848	31.4	369	13	US-10-087-192-1212
27	1547.5	26.3	1170	9	US-09-945-265-2
28	1543.5	26.2	1170	15	US-10-261-164-1
29	1520	25.8	1223	16	US-10-408-765A-295
30	1350.5	23.0	1086	16	US-10-408-765A-1871
31	1229.5	20.9	494	9	US-09-350-259-103
32	1229.5	20.9	494	10	US-09-891-943-103
33	1141.5	19.4	413	9	US-09-350-259-101
34	1141.5	19.4	413	10	US-09-891-943-101
35	1140	19.4	1179	14	US-10-177-550-2
36	1140	19.4	1179	14	US-10-173-551-2
37	1093.5	18.6	1151	10	US-09-884-130-103
38	1093.5	18.6	1151	10	US-09-836-353A-103
39	1093.5	18.6	1179	10	US-09-918-715-250
40	1093.5	18.6	1179	17	US-10-474-794-250
41	1092	18.6	1188	14	US-10-291-265-810
42	1088	18.5	1188	14	US-10-291-265-338
43	1084.5	18.4	1189	10	US-09-884-130-35
44	1084.5	18.4	1189	10	US-09-836-353A-35
45	1084.5	18.4	1189	15	US-10-262-839-4

ALIGNMENTS

RESULT 1

US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902.481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match 100.0%; Score 5884; DB 10; Length 1137;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVCAPQEIIVAAORGSLYQCDYTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVCAPQEIIVAAORGSLYQCDYTGSCPEI 60
QY 61 RLQVPVEAVNMISGLSLAATTSPPQLACGPTVHQTCSNTYVKGLCFPLFGSLNRQOPQK 120
Db 61 RLQVPVEAVNMISGLSLAATTSPPQLACGPTVHQTCSNTYVKGLCFPLFGSLNRQOPQK 120

```
QY 121 PPEALRGCPQSDIAFLIDSGSII PHDPRMKEWSTVMEOLKSKTSLFSLMOYSEEF 180
DB 121 PPEALRGCPQSDIAFLIDSGSII PHDPRMKEWSTVMEOLKSKTSLFSLMOYSEEF 180
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
DB 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSKQELNTVASKPPRDHVFQIN 300
DB 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSKQELNTVASKPPRDHVFQIN 300
QY 301 NFEALKTQNLREKIFAIEGTQTCSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQNLREKIFAIEGTQTCSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTVDSDMDNDAYILGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTVDSDMDNDAYILGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
QY 481 PRGQARWQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
QY 601 PVLRVKAIIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKAIIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPSPIVLRNLF 720
DB 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPSPIVLRNLF 720
QY 721 SLVGTPLSAFNGRLRVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFNGRLRVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
QY 841 SGALKSTCSINHPFFPENSEVTENITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTCSINHPFFPENSEVTENITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVQHQQYQVSNLQSRSLPISLVFLVPV 960
DB 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVQHQQYQVSNLQSRSLPISLVFLVPV 960
QY 961 RLNQTVIWDNRQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 961 RLNQTVIWDNRQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
QY 1021 PFGIIEEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDVSFTLLPGQGAFFVRSQTET 1080
DB 1021 PFGIIEEFNATLKGNSLFDWYIKTSHNHLIIYSTAEILFNDVSFTLLPGQGAFFVRSQTET 1080
QY 1081 KVEPEFVNPLPIVSSVGGLLALLALITAALYKLGFFKRYQKMMSEGGPGABPQ 1137
DB 1081 KVEPEFVNPLPIVSSVGGLLALLALITAALYKLGFFKRYQKMMSEGGPGABPQ 1137
```

RESULT 2

US-09-902-481A-3

; Sequence 3, Application US/09902481A

```
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3
```

```
Query Match 99.5%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQETVAANORGLSYCCDYSTGSCBPI 60
DB 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQETVAANORGLSYCCDYSTGSCBPI 60
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGTVTHQTCSENTYVVKGLCFLFGSNLRQOPK 120
DB 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGTVTHQTCSENTYVVKGLCFLFGSNLRQOPK 120
QY 121 FPEALRGCPQSDIAFLIDSGSII PHDPRMKEWSTVMEOLKSKTSLFSLMOYSEEF 180
DB 121 FPEALRGCPQSDIAFLIDSGSII PHDPRMKELVSTIMEQLKSKTSLFSLMOYSEEF 180
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
DB 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSKQELNTVASKPPRDHVFQIN 300
DB 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSKQELNTVASKPPRDHVFQIN 300
QY 301 NFEALKTQNLREKIFAIEGTQTCSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQNLREKIFAIEGTQTCSSSFHEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTVDSDMDNDAYILGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTVDSDMDNDAYILGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
QY 481 PRGQARWQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
QY 601 PVLRVKAIIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKAIIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPSPIVLRNLF 720
DB 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPSPIVLRNLF 720
```


QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVFLVPV 960
DB 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVFLVPV 960
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
QY 1021 FFGIQQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
DB 1021 FFGIQQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGPPGAEPQ 1137
DB 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGPPGAEPQ 1137

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.5%; Score 5852; DB 9; Length 1153;
Best local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTEAMTFQENARFGQSVWLOGSRVVVVGAPQEIIVANQGRSLYQCDYSTGSCPEI 60
DB 17 FNLDTEAMTFQENARFGQSVWLOGSRVVVVGAPQEIIVANQGRSLYQCDYSTGSCPEI 76
QY 61 RLOQPVAVNMVSLGLSAAATSPQLLACGPTVHOTCENTYVKGHCFLFGSNLRQOPQK 120
DB 77 RLOQPVAVNMVSLGLSAAATSPQLLACGPTVHOTCENTYVKGHCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRRMKFEFVSTYMEQKSKTKLFLSLMOYSEEF 180

DB 137 FPEALRGCPQSDSDIAFLIDGSGSIIPHDFRRMKFEFVSTYMEQKSKTKLFLSLMOYSEEF 196
QY 181 RIHFTTFEFQNNPNRSLIRKIPITQLLORTHTATGLRKVRELFINITNGARKNAKFLFL 240
DB 197 RIHFTTFEFQNNPNRSLIRKIPITQLLORTHTATGLRKVRELFINITNGARKNAKFLVVI 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRVVGVDGAFRSEKSRQELNLTVASKEPRDHVQIN 300
DB 257 TDGEKFGDPLGYEDVPELDREGVIRVVGVDGAFRSEKSRQELNLTVASKEPRDHVQIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQOTGSSSSPEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQOTGSSSSPEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVOSVLGAPRHOHIGLVAMFR 420
DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVOSVLGAPRHOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGVSVCP 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGVSVCP 496
QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYVLF 540
DB 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAYVLF 556
QY 541 HGTSGSGISFSSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISFSSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMBFPREVARNVFECNQDVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 660
DB 617 PVLRVKAIMBFPREVARNVFECNQDVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSTRRTQVGLGTQTCETLKLQCPNCIEDPVSPIVRLNF 720
DB 677 YDLALDSGRPHSRAVFNETKNSTRRTQVGLGTQTCETLKLQCPNCIEDPVSPIVRLNF 736
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVFLVPV 960
DB 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 FFGIQQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
DB 1037 FFGIQQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGPPGAEPQ 1137
DB 1097 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQKDMSEGPPGAEPQ 1153

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy

; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PR
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIVAAQNRSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIVAAQNRSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFGLFSGNLROQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTFKFQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256

QY 241 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 300
Db 257 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVQIN 316

QY 301 NPEALKTIQNLREKI FAIEGTQGTSSSFHEMSQEGFSAAITNGPLLLSTVGSYDWAG 360
Db 317 NPEALKTIQNLREKI FAIEGTQGTSSSFHEMSQEGFSAAITNGPLLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVLAGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVLAGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCLP 480
Db 437 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCLP 496

QY 481 PRGQARWQCDVILGEOGPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 540
Db 497 PRGQARWQCDVILGEOGPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 556

QY 541 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVDLTVCAGQHVLLLRSQ 600
Db 557 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVDLTVCAGQHVLLLRSQ 616

QY 601 PVLVRKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHRAVENETKSTRQTQVLGLTQTCETLKLQIPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHRAVENETKSTRQTQVLGLTQTCETLKLQIPNCIEDPVSPVILRLNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKCKGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRNKTEF 916

QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVMOHQYVSNLQSRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVMOHQYVSNLQSRSLPISLVFLVPV 976

QY 961 RLNOTVIWDRPOVTFSENLSSTCHTKERLPHSHDFLAELRKAPVNVNCISVACQRIQCDIP 1020
Db 977 RLNOTVIWDRPOVTFSENLSSTCHTKERLPHSHDFLAELRKAPVNVNCISVACQRIQCDIP 1036

QY 1021 PFGIOEBENATLKGNSLSDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFFVSOTET 1080
Db 1037 PFGIOEBENATLKGNSLSDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFFVSOTET 1096

QY 1081 KVEPEFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKROYKXMMSEGGPPGABPQ 1137
Db 1097 KVEPEFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKROYKXMMSEGGPPGABPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIVAAQNRSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPEIVAAQNRSLYQCDYSTGSCBPI 76

QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFGLFSGNLROQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKKSKTLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

```
197 RIHFTKEFQNNPNSRLVKPIITQLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPFRSEKSRQELNITVASKPRDHVFOIN 300
257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNITVASKPRDHVFOIN 316
301 NFEALKTIONQUREKIFAIEGTQTGSSSPFHEMSQEGFSAIITNSGPLLTGVSQDWAG 360
317 NFEALKTIONQUREKIFAIEGTQTGSSSPFHEMSQEGFSAIITNSGPLLTGVSQDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 420
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 436
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGQVSVCP 480
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGQVSVCP 496
481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600
557 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616
601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 660
617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 676
661 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRNLF 720
677 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRNLF 736
721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 900
857 SGALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 916
901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQVQSNLQORSLSPLSLVFLVPV 960
917 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQVQSNLQORSLSPLSLVFLVPV 976
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVGSIIVCQRIQCDIP 1020
977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVGSIIVCQRIQCDIP 1036
1021 FFGIOEBFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSTPLPGOGAFVRSOTET 1080
1037 FFGIOEBFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSTPLPGOGAFVRSOTET 1096
1081 KVPEPFVNPPLIVGSSVGLLLALITAAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1137
1097 KVPEPFVNPPLIVGSSVGLLLALITAAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1153
```

RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
```

```
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144, 259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758, 493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221, 950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 ENLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
Db 17 ENLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPPOLACQPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLACQPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 136
Qy 121 FPEALRCPQEDSDIAFLIDGSGSIIIPHDFRRMKWSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRCPQEDSDIAFLIDGSGSIIIPHDFRRMKFVSTVMEQLKKSKTLFSLMOYSEEF 196
Qy 181 RIHFTKEFQNNPNSRLVKPIITQLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 240
Db 197 RIHFTKEFQNNPNSRLVKPIITQLGRTHTATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDEKFGDPLGYEDVPELDREGVIRYVIGVDAPFRSEKSRQELNITVASKPRDHVFOIN 300
Db 257 TDEKFGDPLGYEDVPELDREGVIRYVIGVDAPFRSEKSRQELNITVASKPRDHVFOIN 316
Qy 301 NFEALKTIONQUREKIFAIEGTQTGSSSPFHEMSQEGFSAIITNSGPLLTGVSQDWAG 360
Db 317 NFEALKTIONQUREKIFAIEGTQTGSSSPFHEMSQEGFSAIITNSGPLLTGVSQDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEOTRGQVSVCP 496
Qy 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 660
Db 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRNLF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 840
```

Db 797 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020
Db 977 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGFVRSQDET 1080
Db 1037 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGFVRSQDET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKRYQYKMMSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKRYQYKMMSEGGPPGABPQ 1153

RESULT 7
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 39069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQGSVVQLQGRVVVVGAPQEIIVAAANQGRSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFQGSVVQLQGRVVVVGAPQEIIVAAANQGRSLYQCDYSTGSCPEI 76

Qy 61 RLQVPEAVNNSLGLSLAATSPPOLLAGPTVHOTCSENTYVKGICFLFGSNLRQOPK 120
Db 77 RLQVPEAVNNSLGLSLAATSPPOLLAGPTVHOTCSENTYVKGICFLFGSNLRQOPK 136

Qy 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKKSKTLFLSMQYSEEF 180
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKKSKTLFLSMQYSEEF 196

Qy 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATAGLRKVRELFNITNGARKNAFKLLELL 240
Db 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATAGLRKVRELFNITNGARKNAFKLIVVI 256

Qy 241 TGEKFGDPLGVEDYIPEDLRGVIRYVIGVDGAFRSEKSRQELNVTASKPRDHFQIN 300
Db 257 TGEKFGDPLGVEDYIPEDLRGVIRYVIGVDGAFRSEKSRQELNVTASKPRDHFQIN 316

Qy 301 NFEALKTIONLREKIFAIEGTQTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDMAG 360
Db 317 NFEALKTIONLREKIFAIEGTQTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDMAG 376

Qy 361 GVFLYTSKESKFINMTRVDSMDNDAYLGAAAILLRNVQSILVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKFINMTRVDSMDNDAYLGAAAILLRNVQSILVGLAPRYQHIGLVAMFR 436

RESULT 8

US-09-902-481A-5
; Sequence 5, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

Qy 421 QNTGMBSNANVKCTQICAYFGASLCSVDVDSNGSTDVLILGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMBSNANVKCTQICAYFGASLCSVDVDSNGSTDVLILGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVLYF 540
Db 497 PRGQARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVLYF 556
Qy 541 HGTSGSGLSPSHSORIAGSKLSPRLQYFGQSLSGQDQLTMDGLVDLTVGAGHVLILLRSQ 600
Db 557 HGTSGSGLSPSHSORIAGSKLSPRLQYFGQSLSGQDQLTMDGLVDLTVGAGHVLILLRSQ 616
Qy 601 PVLVKALIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLVKALIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 736
Qy 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVVMVTSYHGVSTKYNFTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 976
Qy 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020
Db 977 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGFVRSQDET 1080
Db 1037 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILLPNDVSFTLLPGQGFVRSQDET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKRYQYKMMSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGILLALITAAALYKLGFFKRYQYKMMSEGGPPGABPQ 1153

US-09-902-481A-5

Query Match 99.4%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 1 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

Qy 61 RLQVPVAVNMISGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSNLRQOPQK 120
Db 61 RLQVPVAVNMISGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSNLRQOPQK 120

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNTNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNTNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 241 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480

Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVOAICAPGEEDNRGAVLY 540
Db 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVOAICAPGEEDNRGAVLY 540

US-09-902-481A-6

Query Match 99.3%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 1 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

Qy 61 RLQVPVAVNMISGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSNLRQOPQK 120
Db 61 RLQVPVAVNMISGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSNLRQOPQK 120

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNTNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPIITQLLGRTHATGLRKVVRELFNTNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 241 TDGEKFGDPLGYEDVPELDREGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480

Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVOAICAPGEEDNRGAVLY 540
Db 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTVOAICAPGEEDNRGAVLY 540

```
Db 481 PRGQRARQCDAVLYGEOGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 600
Db 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 600
Qy 601 PVLVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 601 PVLVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Qy 661 YDLALDSCRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCEIDPVPVLRNLF 720
Db 661 YDLALDSCRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCEIDPVPVLRNLF 720
Qy 721 SLVGTPLSAFGLNRVLAEDAQRFTALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRVLAEDAQRFTALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNTVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Db 781 GPREFNTVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Qy 901 QLELPVKYAVVMVTSKYNFTASENSTRVQVHGYQVSNLQORSLSPLSLVFLVPV 960
Db 901 QLELPVKYAVVMVTSKYNFTASENSTRVQVHGYQVSNLQORSLSPLSLVFLVPV 960
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Qy 1021 PFGIQEEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTEI 1080
Db 1021 PFGIQEEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTEI 1080
Qy 1081 KVEPEVENPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPGABPQ 1137
Db 1081 KVEPEVENPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPGABPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Shimooka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

```
Query Match 99.2%; Score 5836.5; DB 9; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQORGLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVAAANQORGLYQCDYSTGSCEPI 76
```

```
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVYKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVYKGLCFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEWVSTVMEQLKSKSLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKSLFSLMQYSEEF 196
Qy 181 RIHTTFKBFQNNPNRSLIKITQLGRTHATGLRKVVRLEFNITNGARKNAFKILFLL 240
Db 197 RIHTTFKBFQNNPNRSLIKITQLGRTHATGLRKVVRLEFNITNGARKNAFKILVVI 256
Qy 241 TDGKFKGDPGLGYEDVPEADREGVIRVYVIGVDFAFRSEKSEQLNTIASKPRDRHFQVN 300
Db 257 TDGKFKGDPGLGYEDVPEADREGVIRVYVIGVDFAFRSEKSEQLNTIASKPRDRHFQVN 316
Qy 301 NFEALKITQNLREKI FAIECTQTGSSSSPHEMSQEGFSAAITSGNPLLSVGSYDWAG 360
Db 317 NFEALKITQNLREKI FAIECTQTGSSSSPHEMSQEGFSAAITSGNPLLSVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLAVMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLAVMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 496
Qy 481 PRGARARQCDAVLYGEOGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRG-RARQCDAVLYGEOGQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLRSQ 615
Qy 601 PVLVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 616 PVLVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 675
Qy 661 YDLALDSCRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCEIDPVPVLRNLF 720
Db 676 YDLALDSCRPHSRVAFNETKNSRQTQVLGLTQTCETLKLQPNCEIDPVPVLRNLF 735
Qy 721 SLVGTPLSAFGLNRVLAEDAQRFTALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNRVLAEDAQRFTALPPEKKNCGNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNTVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Db 796 GPREFNTVTVVRNDCGDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
Qy 901 QLELPVKYAVVMVTSKYNFTASENSTRVQVHGYQVSNLQORSLSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTSKYNFTASENSTRVQVHGYQVSNLQORSLSPLSLVFLVPV 975
Qy 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1035
Qy 1021 PFGIQEEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTEI 1080
Db 1036 PFGIQEEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTEI 1095
Qy 1081 KVEPEVENPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPGABPQ 1137
Db 1096 KVEPEVENPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKDMWSEGGPGABPQ 1152
```



```
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFFRMKEWSTVMEQLKSKTFLSLMOYSEEP 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGISSRNFMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHTFFKFNQNPNSRIKIPITOLLGTHRTATGLRKVVRELFNITNGARKNAFKILFL 240
Db 198 QHTTFFEFRTSNPLSLASVHQLQGYTYTATQNVVHRLFHASYGAREDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPRDHVFOIN 300
Db 258 TDGKKEGDSLOYKDVIPMADAAGIIRYAIYGVLAFQNRNSWKELNDIASKESQEHFKVE 317
QY 301 NFEALKTIQNLREKIFAIEGTQGTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDQIQNLKEKIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLAVMFR 420
Db 378 GAFLYPPNWSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYYEQTRGGQVSVCP 497
QY 481 PRGQARQCDAVLYGEOGPNGRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRCAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEOGHPWGRFGAALTVLGDVNGDKLTDVIGAPGEBENRCAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAOCHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRYKATMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660
Db 617 PVLWVGVSQMFIPABIPRSAFECEQVQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDIALDGRPHSRVAVNETKNSSTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRNLF 720
Db 677 LDALAPARLSRAIFQETKNRSLSRVRLGLKAHCENFNLLPSCVDSVIPILIRLNF 736
QY 721 SILVGTPLSAFGLNRLPWLAEADAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 TLNGKPLAFNRLPWLALAAQRYFTASLPFEKNCADHICQDNLGIFSFPGLKSLLVG 796
QY 781 GPREFNVTVVNDGEDSVRTQVTFPFLDLSYRKVSTLQNRQSORSWRLACESASTEV 840
Db 797 SNLELNAEVMVWMDGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
QY 841 SGALKSTCSINHPILFPENSEVTNITPDVDSKASIGNKLLKANVTSENMPRNKTEF 900
Db 855 SGTWSTSCRINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRISKITIF 914
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFATAS-ENTSRVMQHQYQVSNLQORSLSIFLFLVP 959
Db 915 QLELPVKYAVYIVVSSHQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974
QY 960 VRLNQTIVDRPQVTFSENLSTCHTKERLPSHDFLAEIRKAPVNVNCSIAVCQIQDI 1019
Db 975 VELNQEAVMVMDEVSHQPNSLRCSSEKIAPPASDFLAHIQNPVLDSCSIAGCLFRCDV 1034
QY 1020 PFGQIEERNATLKNLSFDWKYIKTSHNHLVSTAETILFNDSVFTLLPGQCAFVRSOTE 1079
Db 1035 PSFSQIEELDTLKNLSFGWVRQILQKKVSVSVVAHIFDTSVYSQLPQGBAFNRQITI 1094
QY 1080 TKVEPFEVFNPLPLVIGSVGGLLALITALIYKLGFFPKROYKMMSE 1128
Db 1095 TVLEKYKVHNPPLVIGSIGGLLALITALIYKLVKGFPPKROYKEMBE 1143
```

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

```
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match 58.4%; Score 3436; DB 10; Length 1163;
Best Local Similarity 60.6%; Pred. No. 5.9e-284;
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVGAPQIEIVAAANQRGSLVQCDSYSGCBPI 60
Db 20 FNLDTTELTAFRVDAGFGSDVGVVANSVWVVGAPQKIIAANQIGLGYQCGYSTGACBPI 79
QY 61 PLOYPVEAVNMSGLSLAATTSPPOLLACGCTVHQTCSNTYVVKLCFLFSGNLPRQOQOK 120
Db 80 GLQVPPPEAVNMSGLSLAATTSPPOLLACGCTVHHCGRNNYLTGLCELLEGT--QLTOR 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFFRMKEWSTVMEQLKSKTFLSLMOYSEEF 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGISSRNFMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHTFFKFNQNPNSRIKIPITOLLGTHRTATGLRKVVRELFNITNGARKNAFKILFL 240
Db 198 QHTTFFEFRTSNPLSLASVHQLQGYTYTATQNVVHRLFHASYGAREDAIKILIVI 257
QY 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPRDHVFOIN 300
Db 258 TDGKKEGDSLOYKDVIPMADAAGIIRYAIYGVLAFQNRNSWKELNDIASKESQEHFKVE 317
QY 301 NFEALKTIQNLREKIFAIEGTQGTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360
Db 318 DFDALKDQIQNLKEKIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLAVMFR 420
Db 378 GAFLYPPNWSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 480
Db 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSDTDLVLIGAPHYYEQTRGGQVSVCP 497
QY 481 PRGQARQCDAVLYGEOGPNGRFGAALTVLGDVNGDKLTDVAIGAPCEEDNRCAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEOGHPWGRFGAALTVLGDVNGDKLTDVIGAPGEBENRCAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAOCHVLLRSQ 600
Db 557 HGVLPSPISPSHSQRIAGSKLSSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRYKATMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660
Db 617 PVLWVGVSQMFIPABIPRSAFECEQVQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDIALDGRPHSRVAVNETKNSSTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRNLF 720
```


Db 677 LDALAPGRLSPRAIFOTKMRSLRVRLGLKHAHCENFNLLLSCEVEDSVIPILRLNF 736
Qy 721 SLVGTPLSAGNLAPVLAEDARLFTALFPPEKNCGNNDICQDDLSTFSPMSLDCLVVG 780
Db 737 TLVGKPLLAFLNRLPMLAALAQRYFTASLPPEKNCGADHICQDNLGIGSFPPGLKSLIVG 796
Qy 781 GPREFNVTVVRNDGDSYRQVTFPPPLDLSYKVSFTLQNRQSRWKLACESASSTEV 840
Db 797 SNLENAEVMWNDGDSYGTITFSPAGUSYRYVAEGQKQGLRSLHJTC--CSAPVG 854
Qy 841 SGALKSTCSINHPITFENSESEVTNITFDVDSKASGLNKLKLLKANVTSENNMPTNKTEF 900
Db 855 SQGTWSTSCRNHLIPRGAQITELATFDVSPKAVGLDRLLLIANVSENNIPTSKTIP 914
Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVSNLQORSIPISLVLFP 959
Db 915 QLELPVKYAVYVIVVSHHEQFTKYLNFSSEBESHVAMHRYQVNNLQGRDLFVSNFVWP 974
Qy 960 VRLNQTVIMDRPQVTFSENISSCTKTERLPSSHDFLAELRKAPVNCSTAVCORIQCDI 1019
Db 975 VELNQEAVMWMDVEVSHQNSLRCSSEKIAPPASDFLAHQKNPVLDCSIAQCLRFRCDV 1034
Qy 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
Db 1035 PSFSVQEBDFTLKGNLSFGWVRQILQKYSVSVVAEIIIFDTSVYQLPGQAFMRAQTI 1094
Qy 1080 TKVEPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMWSE 1128
Db 1095 TVLEKYKVNHPILPIVINGSSIGLLLLALITAVLYKVGFFKRYKEMMEE 1143

RESULT 14

US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 57.8%; Score 3401; DB 9; Length 1161;
Best Local Similarity 59.3%; Pred.No.5.8e-281;
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

Qy 1 FNLDENAMTFQENARGFQSVQLOQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 17 FNLDVEPTIFQEDAGGFGQSVQFGSRLVVGAPLEVAANQTRLYDCAATGMCQPI 76
Qy 61 RLQVPVAVNMSLGLSLAATSPQLLACGFTVHQTSCENTYVYKGLCFPLGCSNLRQOPQK 120
Db 77 FLHIRPEAVNMSLGLTAASTNGSRLACGFTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

Qy 121 FPEALRCQPOEDSDIAELIDSGSIIIPHDFRRMKEWSTYMEQLKKSKTLPISIMOYSEEP 180
Db 136 VPDAITPCPCPQHEMDIVELIDSGSIDQDNDFNMKGFFVOAVNGQFEGTDTTUFALMOYSNLL 195
Qy 181 RIHFTFFKEFQNNPNRSLIKPIITQLLGRTHATCLRLRVVRELEFNITNGARKNAFKILFL 240
Db 196 KIHFTTFQFTSPSQSLVDPIVQLKGLTFTATGILTVVTVQLFHHKNGARKSAKKILIVI 255
Qy 241 TDGEKFDPLGYEDVPELDRBGVIRVIVGVDAFRSEKSRQELNVTASPKPRDRHVFQIN 300
Db 256 TDGQKYKDPLEYSVPIPAEKAGIIRYAIVGHAFOQTARQELNLTSSAPPQDHVEKVD 315
Qy 301 NEERALKTIQOLREKIPAIETGTGSSSSSEHEMSQEGFSAATISNGPLISTVCSYDAG 360
Db 316 NFRAALGSIQOLREKIYAVEGTQSRASSSQHEMSQEGFSTALTMDFLFGAVGSFWSG 375
Qy 361 GVFLYTSKEKSTFTNTRVRSDMNDAYLGVAAAIIILNRNVQSLVGLGAPRYQHIGLVAMFR 420
Db 376 GAFLYPENMPTFINMSQENVDMEDSVLGYSTELALWKGVQNLVGLGAPRYQHTGKAVFT 435
Qy 421 QNTGMESNANVKTQIGAYFGASLCSVVDVDSNGSTDVLVIGAPHYYEOTRGGQVSVCP 480
Db 436 QVSRQWRKKAETGCTQIGSYFGASLCSVDVDSGSTDLLIGAPHYYEOTRGGQVSVCP 495
Qy 481 PRGORARWOCDAVLYGEOGPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db 496 PRGORVOMQCDADVLRGEGQHPGRFGAALTVDGVNBEDKLDIDVAIGAPGOENRGAYLYF 555
Qy 541 HGTSGGSISSHQSRLAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVDGAGQHVLLRSQ 600
Db 556 HGASEGSISSHQSRLAGSKLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLRSL 615
Qy 601 PVLRVKAIMEFNPREVARNVPECNDQVVKGEKAGEVVRVCLHVQKSTRDRREBQIQSVVT 660
Db 616 PVLKVGVMARFSPVEAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673
Qy 661 YDIALDSGRPHSRAVFNETKNSTRQVTLGTLCTCETLKLQLPNCIEDPVSPIVLELNF 720
Db 674 FDIALDPGRUTSRAIFNETKNPTLTRKTLGLGHCTETLKLPLDCVDEVDVSPILHLNF 733
Qy 721 SLVGTPLSAGNLAPVLAEDARLFTALFPPEKNCGNNDICQDDLSTFSPMSLDCLVVG 780
Db 734 SLVREPIPSQNLAPVLAEGSODLFTASLPPEKNCQDGLCEGDLGVTLSFGLOITVVG 793
Qy 781 GPREFNVTVVRNDGDSYRQVTFPPPLDLSYKVSFTLQNRQSRWKLACESASSTEV 840
Db 794 SSLELNVITVMNAGDSYGTVVSLSYYPAGLSHRRVSGAQKQPHQSALRACETV-PTED 852
Qy 841 SGALKSTCSINHPITFENSESEVTNITFDVDSKASGLNKLKLLKANVTSENNMPTNKTEF 900
Db 853 EG-LRSRCSVNHPIFHEGSGNGFTIITFDVSYKATLGDRLMLMRASSENKKNKASSKATF 911
Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTAS-ENTSRVMOHQVQVSNLQORSIPISLVLFP 959
Db 912 QLELPVKYAVYVIVVSHHEQFTKYLNFSSEBESHVAMHRYQVNNLQGRDLFVSNFVWP 971
Qy 960 VRLNQTVIMDRPQVTFSENISSCTKTERLPSSHDFLAELRKAPVNCSTAVCORIQCDI 1019
Db 972 VLLNGVAVMDVVMVMEAPSQSL--PCVSEKPKPPQSHSDFLTQISRSPMLDCSIADCLQPRCDV 1029
Qy 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079
Db 1030 PSFSVQEBDFTLKGNLSFGWVRQILQKYSVSVVAEIIIFDTSVYQLPGQAFMRAQME 1089
Qy 1080 TKVEPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMWSE 1128
Db 1090 MVLEEDBVNAIPIMGSSVGCALLLALITATLYKLGFFKRYKEMLEMD 1138

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds

(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5852	99.5	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5852	99.5	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5852	99.5	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5852	99.5	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5852	99.5	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5852	99.5	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5852	99.5	1153	3 US-08-943-363-3	Sequence 3, Appli
8	5852	99.5	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5852	99.5	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5852	99.5	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5821.5	98.9	1152	2 US-08-476-062A-43	Sequence 43, Appl
12	5821.5	98.9	1152	5 PCT-US96-01314-43	Sequence 43, Appl
13	5821.5	98.9	1152	6 5424399-2	Patent No. 5424399
14	3459	58.8	1163	5 US-08-476-062A-44	Sequence 44, Appl
15	3459	58.8	1163	5 PCT-US96-01314-44	Sequence 44, Appl
16	3436	58.4	1163	1 US-08-173-497-4	Sequence 4, Appli
17	3436	58.4	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3436	58.4	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3436	58.4	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3436	58.4	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3436	58.4	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3436	58.4	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3436	58.4	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3436	58.4	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3436	58.4	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3401	57.8	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3401	57.8	1161	1 US-08-286-889-2	Sequence 2, Appli

28	3401	57.8	1161	1 US-08-485-618-2	Sequence 2, Appli
29	3401	57.8	1161	1 US-08-362-652-2	Sequence 2, Appli
30	3401	57.8	1161	2 US-08-605-672-2	Sequence 2, Appli
31	3401	57.8	1161	2 US-08-482-293A-2	Sequence 2, Appli
32	3401	57.8	1161	2 US-08-943-363-2	Sequence 2, Appli
33	3401	57.8	1161	3 US-09-193-043-2	Sequence 2, Appli
34	3401	57.8	1161	4 US-09-688-307A-2	Sequence 2, Appli
35	3401	57.8	1161	4 US-09-350-259-2	Sequence 2, Appli
36	3385.5	57.5	1161	1 US-08-485-618-99	Sequence 99, Appl
37	3385.5	57.5	1161	2 US-08-605-672-99	Sequence 99, Appl
38	3385.5	57.5	1161	2 US-08-482-293A-99	Sequence 99, Appl
39	3385.5	57.5	1161	2 US-08-943-363-99	Sequence 99, Appl
40	3385.5	57.5	1161	3 US-09-193-043-99	Sequence 99, Appl
41	3385.5	57.5	1161	4 US-09-688-307A-99	Sequence 99, Appl
42	3385.5	57.5	1161	4 US-09-350-259-99	Sequence 99, Appl
43	3224.5	54.8	1161	3 US-09-193-043-55	Sequence 55, Appl
44	3224.5	54.8	1161	4 US-09-688-307A-55	Sequence 55, Appl
45	3224.5	54.8	1161	4 US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred.No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVANQRGSLYQCDYSTGSGCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVANQRGSLYQCDYSTGSGCEPI 76

```
QY 61 RLQVPEAVNNMNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVGKLCFLFGSNLRQPOK 120
Db 77 RLQVPEAVNNMNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVGKLCFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWVSTVMEQLKKSKTLFSLMOYSBEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWVSTVMEQLKKSKTLFSLMOYSBEF 196
QY 181 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYVETRGGOVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYVETRGGOVSVCP 496
QY 481 PRGQARWQCDVAVLGEQOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGQARWQCDVAVLGEQOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
QY 541 HCTSGSGTSPSSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAQGHVLLRSQ 600
Db 557 HCTSGSGTSPSSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAQGHVLLRSQ 616
QY 601 PVLVRKAIEMFNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIEMFNPREVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRRTQTVLGLTQTCETILKQLPNCIEDPVPPIVRLNF 720
Db 677 YDLALDGRPHSRVAFNETKNSRRTQTVLGLTQTCETILKQLPNCIEDPVPPIVRLNF 736
QY 721 SLVGTPLSAFGLNRLPVLAEADAQLFTALPPFKKNCNDNIQDDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRLPVLAEADAQLFTALPPFKKNCNDNIQDDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVTVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNSORSQSWRLACESASSTEV 856
QY 841 SCALKSTSCSINHPIFPENSEVTFNIITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SCALKSTSCSINHPIFPENSEVTFNIITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVWQHQQYQVSNLQSRSLPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVWQHQQYQVSNLQSRSLPLSLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPSSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 PFGIQEEFNATLKGNSFDWYIKTSHNHLILVSTAEIILFNDVSFTLLPGQGFVRSQTET 1080
Db 1037 PFGIQEEFNATLKGNSFDWYIKTSHNHLILVSTAEIILFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KYEPEVFPNPLPIVGVSSVGGLLALLALITAALYKLGFFKQYKQMMSEGGPPGABPQ 1137
Db 1097 KYEPEVFPNPLPIVGVSSVGGLLALLALITAALYKLGFFKQYKQMMSEGGPPGABPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNNMNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVGKLCFLFGSNLRQPOK 120
Db 77 RLQVPEAVNNMNSLGLSLAATTSPOLLACGPTVHOTCSNTYVVGKLCFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWVSTVMEQLKKSKTLFSLMOYSBEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEWVSTVMEQLKKSKTLFSLMOYSBEF 196
QY 181 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPDRHVQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
```

```
QY 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAILLRNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAILLRNRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 496
QY 481 PRGORARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db 497 PRGORARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAYLYF 556
QY 541 HGTSGSGISPSHQRINGSKSLRQLYFGQSLSGQDLTMDGLVDLTVGAGQHVLRLRSQ 600
Db 557 HGTSGSGISPSHQRINGSKSLRQLYFGQSLSGQDLTMDGLVDLTVGAGQHVLRLRSQ 616
QY 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSARFNLRPVLAEQAORLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSARFNLRPVLAEQAORLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSORSWRLACASASTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSORSWRLACASASTEV 856
QY 841 SGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMOHQVSNLQORSIPISLFLVPV 976
QY 961 RLQOTVLDWPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLQOTVLDWPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 PFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDVSTLLPGQAFVRSQTET 1080
Db 1037 PFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAELFNDVSTLLPGQAFVRSQTET 1096
QY 1081 KVEPFEVNPFLPLVGSVGGLLLLALITAAALYKLGFKQYKDMSEGGPPGAEPQ 1137
Db 1097 KVEPFEVNPFLPLVGSVGGLLLLALITAAALYKLGFKQYKDMSEGGPPGAEPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3
```

Query Match 99.5%; Score 5852; DB 1; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK 136
QY 121 FPEALRCGPQEDSDIAPLIDGSGSIIPHDFRMRKEWVSTWMEQLKKSKTLFSLMQYSEEF 180
Db 137 FPEALRCGPQEDSDIAPLIDGSGSIIPHDFRMRKEFVSTWMEQLKKSKTLFSLMQYSEEF 196
QY 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGLRKVRVLEPNTNGARKNAFKILFL 240
Db 197 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGLRKVRVLEPNTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPELDREGVIRYVIGVDAFRSEKSRQELNVTASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGVEDVPELDREGVIRYVIGVDAFRSEKSRQELNVTASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQGFSAATNSNGPILLSTVGSYDWAG 360
Db 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQGFSAATNSNGPILLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAILLRNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAAILLRNRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 496
QY 481 PRGORARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAYLYF 540
Db 497 PRGORARWQCDVLYGEGQGWGRFGAALTVDVGVNGDKLTDVAIGAPGEDNRGAYLYF 556
QY 541 HGTSGSGISPSHQRINGSKSLRQLYFGQSLSGQDLTMDGLVDLTVGAGQHVLRLRSQ 600
```

Db 557 HGTSGGSPSHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPVIVRLNF 720
Db 677 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPVIVRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GREENVTVVRNDEGDSVRTQVTFPFLDLVSRKVSTLQNSQRSWRLACESASSTEV 840
Db 797 GREENVTVVRNDEGDSVRTQVTFPFLDLVSRKVSTLQNSQRSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916
QY 901 QLELPVKAYVMVTSYVSHGVSTKYLNFTASENTRVWQHGYQVSNLQORSPLPSLVPLVPV 960
Db 917 QLELPVKAYVMVTSYVSHGVSTKYLNFTASENTRVWQHGYQVSNLQORSPLPSLVPLVPV 976
QY 961 RLNQTWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDDIP 1020
Db 977 RLNQTWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDDIP 1036
QY 1021 FFGIOEFNATLKGNSLSPDWYIKTSHNHLIIYSTAELLFNDVSFTLLPCQAGFVRSQTEF 1080
Db 1037 FFGIOEFNATLKGNSLSPDWYIKTSHNHLIIYSTAELLFNDVSFTLLPCQAGFVRSQTEF 1096
QY 1081 KVEPEVPNPLPIVGVSSVGGLLLLALITALIYKLGFFKQYKDMWSEGGPPGABPQ 1137
Db 1097 KVEPEVPNPLPIVGVSSVGGLLLLALITALIYKLGFFKQYKDMWSEGGPPGABPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32391
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-652-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI VAAQNRGSLYOCDSYSGCEPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEI VAAQNRGSLYOCDSYSGCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 136
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMQKKSKTLFSLMQYSEFP 180
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKEWSTVMQKKSKTLFSLMQYSEFP 196
QY 181 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGRLKVVRELFNITNGARKNAFKILFLL 240
Db 197 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGRLKVVRELFNITNGARKNAFKILFLL 256
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAFRSEKSRQELNNTVASKPRDRHVQFN 300
Db 257 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAFRSEKSRQELNNTVASKPRDRHVQFN 316
QY 301 NFEALKTQNLQRLREKIFALBGTQTCSSSSFEHMSQEGFSAAITNGPLLLSTVGYSDWAG 360
Db 317 NFEALKTQNLQRLREKIFALBGTQTCSSSSFEHMSQEGFSAAITNGPLLLSTVGYSDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNKRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILRNKRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVLYGEOGQPGWRFGAALTULGDVNGDKLTDAI GAPGEENRGAIVLF 540
Db 497 PRGQARWQCDVLYGEOGQPGWRFGAALTULGDVNGDKLTDAI GAPGEENRGAIVLF 556
QY 541 HGTSGGSPSHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGGSPSHSQRAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPVIVRLNF 720
Db 677 YDLALDGRPHSRVAFNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVPVIVRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GREENVTVVRNDEGDSVRTQVTFPFLDLVSRKVSTLQNSQRSWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
Qy 1021 FPGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FPGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1096
Qy 1081 KVPEPEVPNPLPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEPQ 1137
Db 1097 KVPEPEVPNPLPLIVGSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3

; Sequence 3, Application US/08605672

; Patent No. 5817515

; GENERAL INFORMATION:

; APPLICANT: Gallatin, W. Michael

; APPLICANT: Van der Vleren, Monica

; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit

; NUMBER OF SEQUENCES: 103

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 South Wacker Drive, 6300 Sear Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: United States

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/605,672

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/173,497

; FILING DATE: 23-DEC-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/286,889

; FILING DATE: 5-AUG-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/362,652

; FILING DATE: 21-DEC-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659

; REFERENCE/DOCKET NUMBER: 27866/32684

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1153 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-605-672-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSIYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSIYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCFPLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGLCFPLFGSNLRQOPQK 136
Qy 121 FPEALGCCPOEDSDIAFLIDGSGSIIPHDFRRMKEWVSTVMEOLKSKTKLFSLMQYSEEF 180
Db 137 FPEALGCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 196
Qy 181 RIHFTFKFQONNPNRSLIKPIITQLLGRTHATGLARKVRELFNITNGARKNAFKILFL 240
Db 197 RIHFTFKFQONNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILV 256
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVGVGDAFRSEKSRQBELNVTASKPPRHDHFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVGVGDAFRSEKSRQBELNVTASKPPRHDHFQVN 316
Qy 301 NFEALAKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAALTSNGPLLLSTVGSYDWAG 360
Db 317 NFEALAKTIONQREKIFAIEGTQTGSSSFEHMSQEGFSAALTSNGPLLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGQGVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGQGVSVCP 496
Qy 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYL 540
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYL 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTDRDLREGIQSVVT 660
Db 617 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTDRDLREGIQSVVT 676
Qy 661 YDLALDSGRHSRAVENETKNSTRROTGLTCTETLKLQLPNCIEDPVSPVLRLNF 720
Db 677 YDLALDSGRHSRAVENETKNSTRROTGLTCTETLKLQLPNCIEDPVSPVLRLNF 736
Qy 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNIQDDLSITFSFMSLDCLV 780
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNIQDDLSITFSFMSLDCLV 796
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 916
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 FPGIIEFNATLKGNSLDFWYIKTSHNHLIIYSTAIELFNDVSFTLLPGQGFVRSQTET 1080
Db 1037 FPGIIEFNATLKGNSLDFWYIKTSHNHLIIYSTAIELFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KVEPEFVNPLPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
Db 1097 KVEPEFVNPLPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 6

US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENMTFQENARGFGQSVOLQGSRRVVVGAPOEIVAAQNRGSLYQCDSYSGCEPI 60
Db 17 FNLDTENMTFQENARGFGQSVOLQGSRRVVVGAPOEIVAAQNRGSLYQCDSYSGCEPI 76
QY 61 RLQPVVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120
Db 77 RLQPVVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEWSTVMEQLKSKTLFSLMQYSEEF 180

Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFINITNGARKNAFKILFLL 240
Db 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFINITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVDAFRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGEKFGDPLGYEDVPELDRGVIRYVIGVDAFRSEKSRQELNVTASKPPDRHVQIN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNGPLLSVTGSDYDAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSFHEMSQEGFSAAITNGPLLSVTGSDYDAG 376
QY 361 GVFLYTSKEKSTFINMTVDSDMDNAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTVDSDMDNAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASICSVVDVDSNGSTDLVLGAPHYVETGRGOVSVCL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASICSVVDVDSNGSTDLVLGAPHYVETGRGOVSVCL 496
QY 481 PRGQARWQCDVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRCGAVILF 540
Db 497 PRGQARWQCDVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRCGAVILF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLIYQFGQSLSGQDLTMDGLVDLTIVGAQHVLILLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLIYQFGQSLSGQDLTMDGLVDLTIVGAQHVLILLRSQ 616
QY 601 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIEMFNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFEPPLDLSYRKVSTLQNRORSORSLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFEPPLDLSYRKVSTLQNRORSORSLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMVTSHGVSSTKYNFTASENTSRVMQHYOVSNLQBSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSHGVSSTKYNFTASENTSRVMQHYOVSNLQBSLPISLVFLVPV 976
QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORLOCDIP 1020
Db 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORLOCDIP 1036
QY 1021 FFGIOEBFNATLKGNSLDFWYIKTSHNHLIIYSTAIELFNDVSFTLLPGQGFVRSQTET 1080
Db 1037 FFGIOEBFNATLKGNSLDFWYIKTSHNHLIIYSTAIELFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KVEPEFVNPLPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
Db 1097 KVEPEFVNPLPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
DB 17 FNLDENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVAVNMSGLSLAATSPPLLACGPTVHQTCSNTYVYKGLCFGLFSNLRQOPQK 120
DB 77 RLQVPVAVNMSGLSLAATSPPLLACGPTVHQTCSNTYVYKGLCFGLFSNLRQOPQK 136

QY 121 FPEALRGCPQSDSIADLIDSGSIIIPHDPRMKEWYSTVMEQKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQSDSIADLIDSGSIIIPHDPRMKEFVSTVMEQKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPRSLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAKILFL 240
DB 197 RIHFTFKFQNNPRSLVKTITQLLGRTHATGIRKVVRELFINITNGARKNAKILVVI 256

QY 241 TDGKEFGDPLGYEDVIPLEDEGVIRYVIGVDAPRSEKSEQLNTVASKPRDHVFQIN 300
DB 257 TDGKEFGDPLGYEDVIPLEDEGVIRYVIGVDAPRSEKSEQLNTVASKPRDHVFQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNTRVYDSMDNDAYLGVAAAILLRNRVQSLVILGAPRYOHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINNTRVYDSMDNDAYLGVAAAILLRNRVQSLVILGAPRYOHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVILGAPHYEYTRGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVILGAPHYEYTRGQSVCP 496

QY 481 PRQQRARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTDVAIGAPCEEDNRGAVLYF 540
DB 497 PRQQRARWQCDVLYGEOQPGWRFGAALTIVLGDVNGDKLTDVAIGAPCEEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLRVKAIMFNFPREVARNVFECNDQVVRKEAGEVVRVCLHVQKSTRDLREGOIQSVVT 660
DB 617 PVLRVKAIMFNFPREVARNVFECNDQVVRKEAGEVVRVCLHVQKSTRDLREGOIQSVVT 676

QY 661 YDLALDSGRPHSRAVNETKNSTRQTVLGLTQTCETILKQLPNCIEDPVSIVLRLNF 720
DB 677 YDLALDSGRPHSRAVNETKNSTRQTVLGLTQTCETILKQLPNCIEDPVSIVLRLNF 736

QY 721 SLVGTPLSAPGNLRPVLAEDAQRLFTALPPEFKXNCGNDNICODDLSITPSEFMSLDCLVVG 780
DB 737 SLVGTPLSAPGNLRPVLAEDAQRLFTALPPEFKXNCGNDNICODDLSITPSEFMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSQRSLJACESASSTEV 840
DB 797 GPREFNVTVTRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSQRSLJACESASSTEV 856

QY 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANTSENMMPTNKTEF 916

QY 901 QLELPVKYAVYVVTSHGVSTKYLNFTASNTSRVMOHQYOVSNLQORSULPISLVFLVPV 960
DB 917 QLELPVKYAVYVVTSHGVSTKYLNFTASNTSRVMOHQYOVSNLQORSULPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSPHSDFLAELKAPVYVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSPHSDFLAELKAPVYVNCIAVCQRIQCDIP 1036

QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTET 1080
DB 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSOTET 1096

QY 1081 KVEPPEFVNPPLPLIVGSSVGLLLALLITAAALKLGFKKFYKQYKDMSEGGPPGAEPQ 1137
DB 1097 KVEPPEFVNPPLPLIVGSSVGLLLALLITAAALKLGFKKFYKQYKDMSEGGPPGAEPQ 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-193-043-3

Query Match 99.5%; Score 5852; DB 3; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 76

QY 61 RLOVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLRRQPOK 120
DB 77 RLOVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLRRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKEWSTVMBOLKSKTLFSLMOYSBEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKEFVSTVMBOLKSKTLFSLMOYSBEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRRELNITNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRRELNITNGARKNAFKILFLL 256

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSRQELNTVASKPRDHVFQIN 316

QY 301 NFALKTIQNLREKIFAIEGTQTGSSSSEHEMSQGFSAAITSNGLSTVGSYDWAG 360
DB 317 NFALKTIQNLREKIFAIEGTQTGSSSSEHEMSQGFSAAITSNGLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAI LGYAAAII LRNRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRDSDMDNDAI LGYAAAII LRNRVQSLVGLAPRYQHIGLVAMFR 436

QY 421 QNTGWESNANVKTQICAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCP 480
DB 437 QNTGWESNANVKTQICAYGASLCSVDVDSNGSTDLVLIGAPHYETQTRGGQSVCP 496

QY 481 PRGORARQCDAVLYGEGQGWGRFGAALT VLGVDVNGDKLTDVAIGAPGEDNRGAVLYF 540
DB 497 PRGORARQCDAVLYGEGQGWGRFGAALT VLGVDVNGDKLTDVAIGAPGEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLKVKALMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLKVKALMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLENCIEDPSPVTLRLNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLENCIEDPSPVTLRLNF 736

QY 721 SLVGTPTLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPTLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVRNDEGDSYRTQVTFEPFLDLSYKYSTLQNRORSORWKLACESASSTEV 840
DB 797 GPREFNVTVRNDEGDSYRTQVTFEPFLDLSYKYSTLQNRORSORWKLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVYVMTSHGVSTKYNLNTASENTSRVMOHQYQVSNLQGRSLPISLVFLVPV 960

DB 917 QLELPVKYAVYVMTSHGVSTKYNLNTASENTSRVMOHQYQVSNLQGRSLPISLVFLVPV 976

QY 961 RINQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLABLRKAPVNVNCIAVCQRIQCDIP 1020

DB 977 RINQTVIWDPRQVTFSENLSSTCHTKERLPKSHDSFLABLRKAPVNVNCIAVCQRIQCDIP 1036

QY 1021 FPGIOEEFNATLKNLSFDWYIKTSHNLLIVSTAEIILFNDSVFTLLPQOGAFVRSQTBET 1080

DB 1037 FPGIOEEFNATLKNLSFDWYIKTSHNLLIVSTAEIILFNDSVFTLLPQOGAFVRSQTBET 1096

QY 1081 KVEPEVENPLPLIVGSSVGGLLIALITALYKLGFFKROVKOMWSSGGPPGABPO 1137

DB 1097 KVEPEVENPLPLIVGSSVGGLLIALITALYKLGFFKROVKOMWSSGGPPGABPO 1153

RESULT 9
US-09-688-307A-3
; Sequence 3, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-688-307A-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGLYQCDYSTGSCPEI 76

QY 61 RLOVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLRRQPOK 120
DB 77 RLOVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLRRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKEWSTVMBOLKSKTLFSLMOYSBEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKEFVSTVMBOLKSKTLFSLMOYSBEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRRELNITNGARKNAFKILFLL 240
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRRELNITNGARKNAFKILFLL 256

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRYVIGVDAFRSEKSRQELNTVASKPRDHVFQIN 316

QY 301 NFALKTIQNLREKIFAIEGTQTGSSSSEHEMSQGFSAAITSNGLSTVGSYDWAG 360

Db 317 NFEALKTIONQREKIFAIRGTQTGSSSSFEHEMSQEGFSAAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYQTRGGQSVCP 480
Db 437 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYQTRGGQSVCP 496
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYL 540
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYL 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 616
Qy 601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSV 660
Db 617 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSV 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 736
Qy 721 SLVGTPLSFAFNLPRVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDCLV 780
Db 737 SLVGTPLSFAFNLPRVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDCLV 796
Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLISYRKVSTLQNSORSQSWRLACESAS 840
Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLISYRKVSTLQNSORSQSWRLACESAS 856
Qy 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTE 900
Db 857 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTE 916
Qy 901 QLELPVKYAVMVVTSYHGVSTKYNFTASENTSRVMOHQYQVSNLQORSPLISLV 960
Db 917 QLELPVKYAVMVVTSYHGVSTKYNFTASENTSRVMOHQYQVSNLQORSPLISLV 976
Qy 961 RLNOTVLWRDPQVTFSENLSTCHTKERLPSPDPLAELRKAPVNVCSIAVCORIQ 1020
Db 977 RLNOTVLWRDPQVTFSENLSTCHTKERLPSPDPLAELRKAPVNVCSIAVCORIQ 1036
Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVFTLLPGQAFVRSQ 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVFTLLPGQAFVRSQ 1096
Qy 1081 KVEPFEPVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKQDMSEGGPPGAEP 1137
Db 1097 KVEPFEPVNPPLIVGSSVGLLILALITAAALYKLGFFKQYKQDMSEGGPPGAEP 1153

RESULT 10

US-09-350-259-3

; Sequence 3, Application US/09350259

; Patent No. 6620915

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

; APPLICANT: Van der Vieren, Monica

; TITLE OF INVENTION: No. 6620915el Human 2

; FILE REFERENCE: 27866/35004

; CURRENT APPLICATION NUMBER: US/09/350,259

; CURRENT FILING DATE: 1999-07-08

; EARLIER APPLICATION NUMBER: 09/193,043

; EARLIER FILING DATE: 1998-11-16

; EARLIER APPLICATION NUMBER: 08/173,497

; EARLIER FILING DATE: 1993-12-23

; EARLIER APPLICATION NUMBER: 08/286,889

; EARLIER FILING DATE: 1994-08-05

; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21

; EARLIER APPLICATION NUMBER: 08/943,363

; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1153

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-350-259-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAQORGSYQCDYSTGSC 60

Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAQORGSYQCDYSTGSC 76

Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCCSENTYVKGCLFLFGSNLRQ 120

Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCCSENTYVKGCLFLFGSNLRQ 136

Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFPRMKWVSTVMEQLKSKTILFSLMQYSE 180

Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFPRMKWVSTVMEQLKSKTILFSLMQYSE 196

Qy 181 RIHFTKFEFONNPNRSLIKPITQLGRTHATGLRKVRELFNITNGARKNAFKILFL 240

Db 197 RIHFTKFEFONNPNRSLIKPITQLGRTHATGLRKVRELFNITNGARKNAFKILFL 256

Qy 241 TDGEKGDPLGYEDVPELDREGVIRVVGWDAFSEKSEKROELNTVASKPPRDHVFQ 300

Db 257 TDGEKGDPLGYEDVPELDREGVIRVVGWDAFSEKSEKROELNTVASKPPRDHVFQ 316

Qy 301 NFEALKTIONQREKIFAIRGTQTGSSSSFEHEMSQEGFSAAITSNGLPSTVGSYDW 360

Db 317 NFEALKTIONQREKIFAIRGTQTGSSSSFEHEMSQEGFSAAITSNGLPSTVGSYDW 376

Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLV 420

Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLV 436

Qy 421 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYQTRGGQSVCP 480

Db 437 QNTGMESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYQTRGGQSVCP 496

Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYL 540

Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVDVNGDKLTDVAIGAPGEEDNRGAYL 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 600

Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 616

Qy 601 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSV 660

Db 617 PVLRVKAIMBPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSV 676

Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 720

Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLR 736

Qy 721 SLVGTPLSFAFNLPRVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDCLV 780

Db 737 SLVGTPLSFAFNLPRVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDCLV 796

Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLISYRKVSTLQNSORSQSWRLACESAS 840

Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLISYRKVSTLQNSORSQSWRLACESAS 856

Qy 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTE 900

Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVTXYLNFTASENSTRVMQHOYQVSNLQORSLSPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVTXYLNFTASENSTRVMQHOYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
QY 1021 FPGIQEBFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQCGAFVRSQDET 1080
Db 1037 FPGIQEBFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQCGAFVRSQDET 1096
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPO 1137
Db 1097 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPO 1153

RESULT 11

US-08-476-062A-43
; Sequence 43, Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476.062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 98.9%; Score 5821.5; DB 2; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVVKGLCLFLGSLNQRPQPK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVVKGLCLFLGSLNQRPQPK 136
QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKWVSTVMEOLKSKTILSLMOYSEBF 180
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKFVSTVMEQLKSKTILSLMOYSEBF 196
QY 181 RIHFTFKFQNNPNSRLIKPTTOLLGRTHATGLRKVVRELNFINTNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNSRLIKPTTOLLGRTHATGLRKVVRELNFINTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVIPELDRGVIRYVIGVDAPFRSEKSRQELMTVASKPRDRHVQIN 300
Db 257 TDGEKFGDPLGVEDVIPADREGVIRYVIGVDAPFRSEKSRQELMTIASKPRDRHVQIN 316
QY 301 NFEALKTIQNLREKIPIAIEGTOTGSSSSFEHMSOEGFSAAITSGNPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIPIAIEGTOTGSSSSFEHMSOEGFSAAITSGNPLISTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQVSVCLP 480
Db 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLVLI GAPHYEYQTRGGQVSVCLP 496
QY 481 PRGQARWQCDVILYGEQGPWRFGAALTVLGDVNGDKLTDVAIGALGEEDNRGAIVLF 540
Db 497 PRG-RARWQCDVILYGEQGPWRFGAALTVLGDVNGDKLTDVAIGALGEEDNRGAIVLF 555
QY 541 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 615
QY 601 PVLRVKAIIMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 616 PVLRVKAIIMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 735
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNGNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNGNDNICODDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTIVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPREFNVTIVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVTSHGVTXYLNFTASENSTRVMQHOYQVSNLQORSLSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTSHGVTXYLNFTASENSTRVMQHOYQVSNLQORSLSPLSLVFLVPV 975
QY 961 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 976 RLNQTVIWDREPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDDIP 1035
QY 1021 FPGIQEBFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQCGAFVRSQDET 1080
Db 1036 FPGIQEBFNATLKNLSFDWYIKTSHNHLIIIVSTAEIILFNDVSFTLLPQCGAFVRSQDET 1095
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPO 1137

Db 1096 KVEPFEVNPPLVGVSSVGGLLALLIATAALYKLGFFKQYKDMSEGGPPGAEPQ 1152
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US96-01314-43
Query Match 98.9%; Score 5821.5; DB 5; Length 1152;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTFOENARGFGQSVVLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVLOGSRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVVKGLCFGLFSGNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVVKGLCFGLFSGNLRQOPQK 136
Qy 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTLFSLMOYSEEF 196
Qy 181 RIHFTFEFQNNPNSRLIKETIQLGRTHTATGLRKVVRELFNITGARKNAKILFLL 240
Db 197 RIHFTFEFQNNPNSRLIKETIQLGRTHTATGLRKVVRELFNITGARKNAKILFLL 256
Qy 241 TDGEKFGDPLGYEDVIPDLDEGVIRYVIGVDAPFRSEKSRQELNTVASKPRDHVFOIN 300
Db 257 TDGEKFGDPLGYEDVIPDLDEGVIRYVIGVDAPFRSEKSRQELNTVASKPRDHVFOIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSDYDAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSDYDAG 376

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVOSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALIIILNRVOSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVSCPL 480
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVSCPL 496
Qy 481 PRQARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVLYP 540
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVLYP 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 615
Qy 601 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSVVT 660
Db 616 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGIQSVVT 675
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTCTCTLKLQLPNCIEDPVSPIVLINRF 720
Db 676 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTCTCTLKLQLPNCIEDPVSPIVLINRF 735
Qy 721 SLVGTPLSAFGNLRPVLAEQAQLFTALFPPEKNCNDNICODDLSITTFMSLDCLVVG 780
Db 736 SLVGTPLSAFGNLRPVLAEQAQLFTALFPPEKNCNDNICODDLSITTFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSPISLVFLVPV 960
Db 916 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSPISLVFLVPV 975
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
Db 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035
Qy 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
Db 1036 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSQTET 1095
Qy 1081 KVEPFEVNPPLVGVSSVGGLLALLIATAALYKLGFFKQYKDMSEGGPPGAEPQ 1137
Db 1096 KVEPFEVNPPLVGVSSVGGLLALLIATAALYKLGFFKQYKDMSEGGPPGAEPQ 1152
RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO:2;
; LENGTH: 1152
; 5424399-2
Query Match 98.9%; Score 5821.5; DB 6; Length 1152;

Best Local Similarity 98.9%; Pred. No. 0;			
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;			
QY	1	FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPQEIIVAAQNORGLYQCDYSTGSCBPI	60
DB	17	FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPQEIIVAAQNORGLYQCDYSTGSCBPI	76
QY	61	RLQVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	120
DB	77	RLQVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	136
QY	121	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKSKTKLFLSMOYSEEF	180
DB	137	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKSKTKLFLSMOYSEEF	196
QY	181	RIHFTFKERQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240
DB	197	RIHFTFKERQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI	256
QY	241	TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFRSEKSRQBLNTVASKPPRDHVFQIN	300
DB	257	TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFRSEKSRQBLNTVASKPPRDHVFQIN	316
QY	301	NFEALKTIONQREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDMAG	360
DB	317	NFEALKTIONQREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDMAG	376
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAATILNRVQSLVILGAPRYQHIGLVAMPR	420
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAATILNRVQSLVILGAPRYQHIGLVAMPR	436
QY	421	QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP	480
DB	437	QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP	496
QY	481	PRGQARWOCDAVLYGEGQPMGRFGAALTIVLGDVNGDKLTDAI GAPGEEDNRGAVLYF	540
DB	497	PRG-QARWOCDAVLYGEGQPMGRFGAALTIVLGDVNGDKLTDAI GAPGEEDNRGAVLYF	555
QY	541	HGTSGSGTSPSHSQRAGSKLSPRQYFGQSLSGQDLTMDGLVDLTVGAQGHVILLRSQ	600
DB	556	HGTSGSGTSPSHSQRAGSKLSPRQYFGQSLSGQDLTMDGLVDLTVGAQGHVILLRSQ	615
QY	601	PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
DB	616	PVLRVKAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	675
QY	661	YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF	720
DB	676	YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF	735
QY	721	SLVGTPLSAFGLNLRPLVAEDAQRLETFALPPEKNCNDNICODDLSITPFSMSLCLVVG	780
DB	736	SLVGTPLSAFGLNLRPLVAEDAQRLETFALPPEKNCNDNICODDLSITPFSMSLCLVVG	795
QY	781	GPREFNVTVVRNDGEDSVRTQVTFPPFDLDSYRKVSTLQNSQRSWRLACESASSTEV	840
DB	796	GPREFNVTVVRNDGEDSVRTQVTFPPFDLDSYRKVSTLQNSQRSWRLACESASSTEV	855
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSNNPRNKTEF	900
DB	856	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSNNPRNKTEF	915
QY	901	QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVWQHOYQVSNLQORSLPLSLVFLVPV	960
DB	916	QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVWQHOYQVSNLQORSLPLSLVFLVPV	975
QY	961	RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP	1020
DB	976	RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP	1035
QY	1021	PFGIQEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAAILFNDSVFTLLPQOGAFVRSQSTET	1080

DB	1036	PFGIQEEFNATLKGNSLSPDWYIKTSHNHLIIYSTAAILFNDSVFTLLPQOGAFVRSQSTET	1095
QY	1081	KVEPEVENPULIIVGSSVGGILLALITAAALYKLGFKROVKDMMSEGGPGABPQ	1137
DB	1096	KVEPEVENPULIIVGSSVGGILLALITAAALYKLGFKROVKDMMSEGGPGABPQ	1152
RESULT 14			
US-08-476-062A-44			
; Sequence 44, Application US/08476062A			
; Patent No. 5877275			
; APPLICANT: Arnaout, M. Amin			
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY			
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS			
; NUMBER OF SEQUENCES: 53			
; CORRESPONDENCE ADDRESS:			
; ADDRESS: Fish & Richardson P.C.			
; STREET: 225 Franklin Street			
; CITY: Boston			
; STATE: MA			
; COUNTRY: US			
; ZIP: 02110-2804			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Diskette			
; COMPUTER: IBM Compatible			
; OPERATING SYSTEM: Windows95			
; SOFTWARE: FastSEQ for Windows Version 2.0			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/476,062A			
; FILING DATE: 07-JUN-1995			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: 08/216,081			
; FILING DATE: 21-MAR-1994			
; APPLICATION NUMBER: 07/637,830			
; FILING DATE: 04-JAN-1991			
; APPLICATION NUMBER: 07/539,842			
; FILING DATE: 18-JUN-1990			
; APPLICATION NUMBER: 07/212,573			
; FILING DATE: 28-JUN-1988			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Freeman, John W.			
; REGISTRATION NUMBER: 29,066			
; REFERENCE/DOCKET NUMBER: 00786/068003			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 617/542-5070			
; TELEFAX: 617/542-8906			
; TELEX: 200154			
; INFORMATION FOR SEQ ID NO: 44:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1163 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-476-062A-44			
Query Match 58.8%; Score 3459; DB 2; Length 1163;			
Best Local Similarity 60.9%; Pred. No. 1.2e-289;			
Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;			
QY	1	FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPQEIIVAAQNORGLYQCDYSTGSCBPI	60
DB	20	FNLDTENAMTFQENARGFQGSVVQLOGSRVVVGAPQEIIVAAQNORGLYQCDYSTGSCBPI	79
QY	61	RLQVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	120
DB	80	GLQVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	137
QY	121	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKSKTKLFLSMOYSEEF	180
DB	138	LPVSQECRQEQDQIVLIDGSGSIISRNFMFVRAVISOQFORPSTQFSLMQSNKF	197
QY	181	RIHFTFKERQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481b-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPGGAEFQ 1137
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	1 RWHU1B	cell surface glyco
2	4470	76.1	1153	2 S00551	leukocyte surface
3	3479	59.2	1163	1 RWHU1C	cell surface glyco
4	1533.5	26.4	1170	2 S03308	cell surface glyco
5	1538.5	26.2	1163	2 I56126	lymphocyte fuction
6	1153	19.6	1179	2 A53213	integrin alpha-E c
7	1103.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1085	18.5	1170	2 I45914	integrin alpha 2 s
9	1071	18.2	1178	2 S44142	VLA-2 protein homo
10	1068	18.2	1181	2 A33998	integrin alpha-2 c
11	1062	18.1	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	635	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 JC7294	alphaP integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	542.5	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	integrin alpha-v c
23	532	9.1	1073	2 B36429	fibronectin recept
24	531.5	9.0	1072	2 A38457	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-6 c
26	526	9.0	1048	2 A27421	integrin alpha-3 c
27	525.5	8.9	1091	2 A41543	integrin alpha-5 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit

ALIGNMENTS

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N: Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C: Species: Homo sapiens (man)
C: Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004
R: Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.
J. Biol. Chem. 263, 12403-12411, 1988
A: Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CR

B.
A: Reference number: A31108; MUID: 88315033; PMID: 2457584
A: Accession: A31108
A: Molecule type: mRNA

A: Residues: 1-1153 <COR>

A: Cross-references: UNIPROT: P11215; GB: J03925; NID: g187284; PIDN: AAA59544.1; PID: g307148

A: Note: part of this sequence was confirmed by protein sequencing

R: Arnaut, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.
J. Cell Biol. 106, 2153-2158, 1988

A: Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M

A: Reference number: A28915; MUID: 88257215; PMID: 2454931

A: Accession: A28915

A: Molecule type: mRNA

A: Residues: 1-499, 501-965, 'P', 967-1153 <ARN>

A: Cross-references: GB: M18044; GB: J03270; GB: M19664; GB: X07421; NID: g186935; PIDN: AAA594

A: Note: the authors translated the codon TAC for residue 1129 as Thr

A: Note: part of this sequence, including the amino end of the mature protein, was confir

R: Shelley, C.S.; Arnaut, M.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 10525-10529, 1991

A: Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A: Reference number: A41600; MUID: 92073318; PMID: 1683702

A: Accession: A41600

A: Molecule type: DNA

A: Residues: 1-9 <SHE>

A: Cross-references: GB: M76724; NID: g180018; PIDN: AAA58410.1; PID: g553215

R: Arnaut, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.
Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A: Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi

A: Reference number: A94193; MUID: 88190151; PMID: 2833753

A: Accession: A30892

A: Molecule type: mRNA

A: Residues: 917-1042 <AR2>

A: Cross-references: GB: M18044

R: Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A: Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A: Reference number: A32218; MUID: 89098893; PMID: 2563162

A: Accession: A32218

A: Molecule type: mRNA

A: Residues: 9-1153 <HIC>

A: Cross-references: GB: J04145; NID: g189068; PIDN: AAA59903.1; PID: g386975

A;Note: part of this sequence was confirmed by protein sequencing
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
J. Immunol. 150, 480-490, 1993
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
n during evolution.
A;Reference number: A46526; MUID:93123748; PMID:8419480
A;Accession: A46526
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-499,501-1153 <FILE>
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA824821.1; PID:G263049
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
Biochim. Biophys. Acta 874, 368-371, 1986
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
A;Reference number: A90664; MUID:87076671; PMID:3539202
A;Accession: A26091
A;Molecule type: protein
A;Residues: 17-31 <PIE>
A;Experimental source: granulocytes
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
Blood 79, 865-870, 1992
A;Title: Characterization of the myeloid-specific CD11b promoter.
A;Reference number: I52567; MUID:12144986; PMID:1346576
A;Accession: I52567
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-9 <RES>
A;Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
C;Genetics:
A;Gene: GDB:ITGAM; CR3A
A;Cross-references: GDB:120599; OMIM:120980
A;Map position: 16p11.2-16p11.2
A;Note: promoter contains a GATA motif and two Spl consensus binding sites
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
F;17-1108/Domain: extracellular #status predicted <EXT>
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;465-473/Region: calcium/magnesium binding #status predicted
F;530-538/Region: calcium/magnesium binding #status predicted
F;533-601/Region: calcium/magnesium binding #status predicted
F;1109-1134/Domain: transmembrane #status predicted <TM>
F;1135-1153/Domain: intracellular #status predicted <INT>
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGSVVQLQGSRVVVGAPQBIIVANQSGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGSVVQLQGSRVVVGAPQBIIVANQSGLYQCDYSTGSCPEI 76

Qy 61 RLQVPVEAVNMSGLSLAATSPPLLACGTVHTQTSNTYVVKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAATSPPLLACGTVHTQTSNTYVVKGLCFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDIAFLVDGSGSIIIPHDFRRAKEFTSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLVDGSGSIIIPHDFRRAKEFTSTVMEQLKSKTLFSLMOYSEEF 196

Qy 181 RIHFTTFKFNPNPRSLIKPITQLGRTHATGIRKVVRLFNITNGARQNAFKILILI 240
Db 197 RIHFTTFKFNPNPRSLIKPITQLGRTHATGIRKVVRLFNITNGARQNAFKILIVI 256

Qy 241 TDGEKFGPLGYEIVIPADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFOIN 300
Db 257 TDGEKFGPLGYEIVIPADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFOVN 316

Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILLSTVGSVDWAG 360
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILLSTVGSVDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAIILNRVQSLVLAIGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAIILNRVQSLVLAIGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASCLSDVDNSNGSTDVLIGAPHYYEITRGQGVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASCLSDVDNSNGSTDVLIGAPHYYEITRGQGVSVCP 496

Qy 481 PRQQRARWCDAVLYGEOQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEDNRGAVLYF 540
Db 497 PRQQRARWCDAVLYGEOQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGGAQGHVLLRSQ 616

Qy 601 PVLRVKAIMFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676

Qy 661 YDLALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRNF 736

Qy 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALPFPEKNCNDNICODDLSITTFMSLDCLVVG 780
Db 737 SLVGTPLSAPGNLRPVLAEADAQRLFTALPFPEKNCNDNICODDLSITTFMSLDCLVVG 796

Qy 781 GPREENVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRSWRLACSSASSTEV 840
Db 797 GPREENVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRSWRLACSSASSTEV 856

Qy 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSGSLPISLVFLVPV 960
Db 917 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVSNLQSGSLPISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1036

Qy 1021 FPGIQEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1080
Db 1037 FPGIQEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1096

Qy 1081 KVBPPEFVNPPLITVGVSSVGLLLALITAAALYKGFQKQYKDMSEGPPGAEPO 1137
Db 1097 KVBPPEFVNPPLITVGVSSVGLLLALITAAALYKGFQKQYKDMSEGPPGAEPO 1153

RESULT 2

S00551

Leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C;Accession: S00551; I59078

R;Pyrula, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

A;Molecule type: DNA

A;Residues: 1-1153 <PVT>

A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G5298

A;Note: the authors translated the codon CAC for residue 569 as Glu

R;Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor
A;Reference number: I59078; MUID:86287312; PMID:2942940
A;Accession: I59078
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 11-44 <RES>
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C;Genetics:
A;Gene: Mac-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homodimer; cell adhesion; glycoprotein; transmembrane protein
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;
Best Local Similarity 73.9%; Pred. No. 7.9e+303;
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

QY 1 ENLDTENAMTFQENARFGQSVVOLGSRVVGAPQEIIVAANORGSLYQCYDSTGSCPT 60
Db 17 ENLDTENAMTFQENARFGQSVVOLGSRVVGAPQEIIVAANORGSLYQCYDSTGSCPT 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSTENTYKGLCFGLGSLRQOPQ 120
Db 77 PLQVPVEAVNMSLGLSLAVSTVPOQLLACGPTVHQCSTENTYKGLCFGLGSLRQOPQ 136

QY 121 FPEALRGCPQSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTLFSLMOYSEEF 196

QY 181 RHFTFKFQNNPRLKIPITOLLGRTHATGIRKVRRLPNYNGARKNAFKILILI 240
Db 197 RHFTFKFQNNPRLKIPITOLLGRTHATGIRKVRRLPNYNGARKNAFKILILI 256

QY 241 TDGEKFGDPLGYEDVPEADREGRVYVIGVDGAPRSEKSRORLNTVASKPRDHFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGRVYVIGVDGAPRSEKSRORLNTVASKPRDHFQIN 316

QY 301 NFEALKTIONLREKIFAIEGTQTSSTSSFEHMSQEGFSAITSNGLPSTVGSVDWAG 360
Db 317 NFEALKTIONLREKIFAIEGTQTSSTSSFEHMSQEGFSAITSNGLPSTVGSVDWAG 376

QY 361 GVFLYTSKEKSTFNNTRVDSMDNAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFNNTRVDSMDNAYLGAAIILRNVRQSLVIGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDLVIGAPHYETRGQSVVCP 480
Db 437 ENFGTWEHTSIKSGQISYFGASLCSVDVDSNGSTDLVIGAPHYETRGQSVVCP 496

QY 481 PRGQARWQCDVAVLGEQGPWRGAALTVLGDVNGDKLTDVAIGAPGEQENQGVNIF 540
Db 497 PRGQARWQCDVAVLGEQGPWRGAALTVLGDVNGDKLTDVAIGAPGEQENQGVNIF 555

QY 541 HCTSGSGISPHSQRISAGSKLSPRIQYFGQSLGSGQDLTMDGLVDLTVCAQGHVLLRSQ 600
Db 556 YGASTASLASASHRII GAHFSFGIYFGQSLGSGQDLTMDGLVDLTVCAQGHVLLRSQ 615

QY 601 PVLVRKALMEFNPVARNVFCNDQVVKGEKAGEVRVCLHVOKSTORLREGQIOSVVT 660
Db 616 PVLVRKALMEFNPVARNVFCNDQVVKGEKAGEVRVCLHVOKSTORLREGQIOSVVT 675

QY 661 YDLALDSGRPHSVAFNETPNSRRQTQVLGLTQTCETLKLPLNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSVAFNETPNSRRQTQVLGLTQTCETLKLPLNCIEDPVPVILRLNF 735

QY 721 SILVGTPLSAGNLRVLAEDAQRLLTALFPPEKNGCNNDICODDLISITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLNLRVLAEDAQRLLTALFPPEKNGCNNDICODDLISITFSFMSLDCLVVG 795

QY 781 GPREFNVTVTRNDGEDSVTRQVTFPPFDLSYKRVKSTLQNRQSRWRL-ACESASSTE 839
Db 796 GPQDFNMSVTRNDGEDSVTRQVTFPPFDLSYKRVKSTLQNRQSRWRL-ACESASSTE 855

QY 840 VSGALKSTGSIINHPFPPENSEVTENITFDVDSKASIGKLLIKANVTSENNPRINKTE 899
Db 856 GHGALKSTGSIINHPFPPENSEVTENITFDVDSKASIGKLLIKANVTSENNPRINKTE 915

QY 900 FQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVQHQYQVSNLQSRSLPISLVFLVP 959
Db 916 FQLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVQHQYQVSNLQSRSLPISLVFLVP 975

QY 960 VRLNQTVINDRQVTFSENLSTSTCHTKERLPHSDFLAEARVAVVNCISIAVCQRIQCDI 1019
Db 976 VQINNVTVMDHPQVIFSQNLSSACHTEQKSPHNSPRDQLERTPVLNCSVAVCKRIQCDL 1035

QY 1020 PFFGIOEFENATLKGNSLFDWYIKTSHHLLIVSTAEILFNDSVFTLLPQOGAFVRSQTE 1079
Db 1036 PFFGIOEFENATLKGNSLFDWYIKTSHHLLIVSTAEILFNDSVFTLLPQOGAFVRSQTE 1095

QY 1080 TKVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1096 TKVEPEVNPPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 3
RWHLIC

N;Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A35543; A35543; S00864
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A;Reference number: A35543
A;Contents: erratum
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-1163 <COR>
A;Cross-references: UNIPROT:P20702
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Accession: A35543; MUID:90153906; PMID:2303426
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A3584
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755, 'L', 757-1163 <CO3>
A;Cross-references: GB:M61695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homodimer; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;1108-1133/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61.89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.2%; Score 3479; DB 1; Length 1163;
Best Local Similarity 61.1%; Pred. No. 9.3e-234;
Matches 690; Conservative 141; Mismatches 292; Indels 6; Gaps 4;

QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
DB 20 ENLDTEBELTAFRVDSAGFGSVVQYANWSVVVVGAPQKITANQOTGGLYQCYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFGLFGSNLRQOPQK 120
DB 80 GLQVPPEAVNMSLGLSLASTTSPQLLACGPTVHHEGGRNMYLTGLCFLLGPT--QLTOR 137
QY 121 FPBALRGCPQEDSDIAFLVQSGSIIIPDFRRAKEFTTMBQELKSKTILFSLMQSEEF 190
DB 138 LPVSROCEPQEQDVLIDGSGSISRNFAFMNFVRAVISQFQRPSTQFSLMQFSNKF 197
QY 181 RIHFTKFEQNNPNSRLIKPITOLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240
DB 198 QTHFTFEFRRTSNPLSLLASVHQLQGFTYTATAIQNVHRLPHASVGARDATKILIVI 257
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNVTVASKPPRDHVFQIN 300
DB 258 TDGKKEGSLDYKDVIPMAAAGIIRYAIGVGLAFQNRNGWKELNDIASKPSQEHIEKVE 317
QY 301 NFEALKTIONLREKIFAIEGTQTGSSSSPEHEMSQGFSAATISNGPILLSYVGSYDWAG 360
DB 318 DFDALKDIONLREKIFAIEGTQTSSTSSFELEMAQEGFAVFDPDGVILGAVGVSFTWSG 377
QY 361 GVFLYTSKESKSTFINMTFRVDSMDMDAYLGAALILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINMSQENVMDSDVLGYSTELALMKGVQSLVGLGAPRYQHTGKAVIFT 437
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYQEYTRGGQVSCVPL 480
DB 438 QVSRQWRMKAETVTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYQEYTRGGQVSCVPL 497
QY 481 PRGORARWODAVLYGQGGQWGRGALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF 540
DB 498 PRGWR--RWCDAVLYGEGQWGRGALTVLGVNGDKLTDVVGIVGAPGEENKGAVLYF 556
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGVLPSPISPSHSORIASGLSRLQYFGQSLGGQDLTQGLVDLTVGARGQVLLRTR 616
QY 601 PVLRVKAIMERNPREVARNFECDNQVVKGEAGEVRVCLHVQKSRDRRLREGQIQSVVT 660
DB 617 PVLWVGVSQMFPAPBIPRPAPECREQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSRVAFVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPSYPIVLRNF 720
DB 677 LDALDPGLSLPRATFQTKRSLSRVRVLGLKAHCENFNLLPSCVEDSVTPITLRNF 736
QY 721 SLVGTPLSAFNLAPVLAEDAQRLLFTALFPPEKXGNDNIQDDLSITFSFMSLCLVVG 780
DB 737 TLVGPFLAFLNPLMLAADAQRYFTASLPPEKXGADHIQDNLGISFSFPLGKSLVVG 796
QY 781 GPRFNVTVVRNDEGSYRTQVTFPFLDLISYRKVSTLQNRORSORSLACESASSTEV 840
DB 797 SNLEINAEVWVWVNDGDSYGTITTFSPHAGLSYRYVAEGQKQGLRSLHLTCDSPVG-- 854
QY 841 SGALKSTSCSNHPTFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPNTRNKTEF 900
DB 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVLGDRLLLLTANVSSENTPRTSKTTF 914
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFNTAS--ENTSRVMOHQYVNSLQORSIPISLVFLVP 959
DB 915 QLELPVKYAVYVVSHEQFTKYLNFSSSEKEKSHVAMHRYQVNNLQORDLPVSNFWVP 974
QY 960 VELNQTVIWDPRQVTFSENLSSTCHTKERLPHSHDFLAEKAPVNVCSIAVCORIQDI 1019
DB 975 VELNQAEVWMDVEVSHPPNPSLRCSSEKIAPSPADFLAHIQKVPVLDSCSGLCFRCDV 1034

QY 1020 PFGIQEENFNAIKGNLSFDFWIKTSHNHLIIVSTAEILFNDSVFTLLPFGQGFVRSQTE 1079
DB 1035 PFSVQGEELDFILKGNLSFGWVRQIIQKKVSVVVAEITFTSVYSQLPQGEAFMRAQTT 1094
QY 1080 TKVEPPEVNPUPLIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSE 1128
DB 1095 TVLEKYKHVNPPTLVGSSSTGGLLLLALITAVLYKVGFFKRYQKEMEE 1143

RESULT 4
S03308
cell surface glycoprotein CD11a precursor - human
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text_change 09-Jul-2004
C;Accession: S03308; A47458; A47565; A48759; S36044
J;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.
J;Cell Biol. 108, 703-712, 1989
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit
A;Reference number: S03308; MUID:89139587; PMID:2537322
A;Accession: S03308
A;Molecule type: mRNA
A;Residues: 1-1170 <LAR>
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CRA68
A;Note: part of this sequence was confirmed by protein sequencing
R;Cornwell, R.D.; Gollabon, K.A.; Hickstein, D.D.
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pro
A;Reference number: A47458; MUID:93248261; PMID:8097887
A;Accession: A47458
A;Molecule type: DNA
A;Residues: 1-20 <COR>
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993
A;Title: Identification of cell-specific and developmentally regulated nuclear factors
A;Reference number: A47565; MUID:93281759; PMID:8099450
A;Accession: A47565
A;Molecule type: DNA
A;Residues: 1-20 <SHE>
A;Cross-references: GB:M95609
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.
J. Biol. Chem. 269, 19305-19311, 1993
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter.
A;Reference number: A48759; MUID:93374910; PMID:8103515
A;Accession: A48759
A;Molecule type: DNA
A;Residues: 1-20 <NUB>
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406
C;Genetics:
A;Gene: GDB:ITGAL; CD11A
A;Cross-references: GDB:119757; OMIM:153370
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr
P;1-25/Domain: signal sequence #status predicted <SIG>
P;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted
P;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.4%; Score 1553.5; DB 2; Length 1170;
Best Local Similarity 34.7%; Pred. No. 1.5e-99;
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAQNGSLYQCDYSTGSC 58
DB 26 YNLDVRGARSFSPRAGRHFGYRVLQV--GNGVIVGAPGE---GNSTGLYQCQSGTGCHL 81
QY 59 PIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFGLFGSNLR--- 115
DB 82 PVTLR--GSNYTSKYLGMTLATDPTDGSILACDPGLSRCTCDQNTYLSGLCYLFRQLQGP 140
QY 116 --QPOKPFPEARLGCPEQSDSDIAFLVDGSGSIIIPHDFRAKSFISTVMBQLKSKTLFSLM 174

Db 141 LQGRGFGQECIKG-----NVDLVFLFDGMSLQDFQKILDFMKDVMKKLSNTSYQFAAV 196
QY 175 YSEFEFRIHFTKPEQNPNPRLSIKPIQTOLGRTHATGIRKVVRELFNITNGARKNAP 234
Db 197 QFSTSYKTEFDSDVVKWKDPALLKHVKHMLLINTGAINVATEVFEELGARPDAT 256
QY 235 KILLITDGEKFGDPLGVEDVPEADRGVIRYVIGVDAPRSEKSRQELNIVASKPPRD 294
Db 257 KVLIIITDGE--ATDSGNIADAKD-----IIRYIIIGIKGHFQTKESQETLHKFASKPASE 309
QY 295 HVFOINNEALKTIQNLREKIPATEGTQTSSSSFEHMSQEGFSAITNGPILLSVVG 354
Db 310 FVKILDTFEKLDLTELQKKIYIIEGTSKODLTSFNMLSSSGISADLSRGHVVGAAGV 369
QY 355 SYDWAGGVF-LYTSKEKSTFINMTRVDSMDMDAYLYGAAA-IILNRVQSLVLGAPRYOH 412
Db 370 AKDWAGGFLDKADLQDDTFIGNPLETPVRAVILGYVTWLPBROKTSILASGAPRYOH 429
QY 413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQT 470
Db 430 MGRVLLFQEPQGGHWSQVQTHGTQIGSYFGGELCGVDVDQDGETELLIGALPLFYGEQ 489
QY 471 RGGQVSCPLPRGORARWQCDVAV--LYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGAP 528
Db 490 RGRVFYI-----QRRQLGFEVSELODGPYPIGRFGEAITALTLDINGDGLVDVAVGAP 544
QY 529 GEEDNRGAVLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTV 588
Db 545 LEE--QGANVIFNGRHG-GLSPQSPQRLEGTVGLSGIOWFGRSIHGVKDLGGDLADVAV 601
QY 589 GAQGHVLLLRSQPVLRKALMEFNPREVARNVFECNDQV--KGKEAGEVRVCLHVQKSTR 647
Db 602 GABSQMTVLSSRPVDMVTILMSFPABIPVHEVECSYSTSNKMEGVNITICQI-KSLY 660
QY 648 DRLREGIQSVVTVLDALDSGRPHSRAVENETKNSITRQTVGLTGTCTELKLQLPNCI 707
Db 661 PQF-QGELVANILYTLQDGHTRRRGLFGGRHELRNIAVT-TSMSCDTDFSFHFPCV 718
QY 708 EDVPSPIVLNPLNSL---VGTPLS--AFGN-----LRPVLAEDAQRFLTALFFPEKNCGN 757
Db 719 QDLISPINVSLNFSLMEEGTPRDQRAQKDIPIPLRPSLHSETWEI-----PFENKCGE 773
QY 758 DNICQDDLSITFMSLDCLVGGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVS 817
Db 774 DKKEANLRVSPARSRALRLTAFASLSVELSLSLEADAYVWQDLDFPPGGLSRKVE 833
QY 818 TLQNRQSQRSWRLACES--ASSTEVSGALKSTCSINHPIFPENSEVTFNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCPELPEERLLSRAL---SCNVSSPIFKAGHSVALQMMENTLVNS 887
QY 876 LGNKLKLLKANVTSENN---MPTNKTEFQLELPVKYAVVMVVTSHGVSTKYLNFTASEN 931
Db 888 WGDSDVELHANVTENNEDSDLLDENSAITTI--IPILYPINILIQODESTLVSFTPKGP 944
QY 932 TSVMOHQYQV---SNLQORSPL-ISLVLVPLVRLNQTVWDRPQVTFNSENLSSTCHTK- 986
Db 945 KTHQVKHMYQVRIQPSIHQHNITPLEAVVGVPOPPSEGPITHQWSYOMBPV--PCHYED 1002
QY 987 -BRLPSSHSD--FLAELRKAPVNCSTAVCQRIQCDDIPFGIGQEEFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAABPCPLGALFRCPVW-----FRQELIVQVIGTLELVGEIE 1044
QY 1044 TSHNHLIVSTABIIFNDSVFTLLPQGAFAVRSQTEKVEPEVFNPLPLIVGSSVGLL 1103
Db 1045 AS-SMFSLCSLSISFNSSKHFLYGSNASL-AQVVMKVDDVYVEKQMLYLYVLSGIGLL 1102
QY 1104 LLALITAAIYLGFFKQVQKQVMMSEG-GPPGAP 1136
Db 1103 LLLLIFIVLYKGVFFKRNLRKERWAGRGVNGIP 1136

RESULT 5
156126

lymphocyte fuction-associated molecule-1-alpha - mouse
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I56126
R:Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A:Reference number: I56126; MUID:91268576; PMID:2051027
A:Accession: I56126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1163 <RES>
A:Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39436.1; PID:g19878
C:Genetics:
A:Gene: LFA-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F:151-315/Domain: von Willebrand factor type A repeat homology <FWAL>

Query Match 26.2%; Score 1538.5; DB 2; Length 1163;
Best Local Similarity 34.2%; Pred. No. 1.6e-98;
Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;
QY 1 FNLDTENAMTFQENA-RGFGQSVVOLQSGRVVVGAPQBIIVAANQSGSLYQCDYSTGCEP 59
Db 24 YNLDTRTQTSLAQAGRHFGYVLQIEDG-VVVGAPGE---GDNVTGLYHCHRTSSEFCQP 79
QY 60 IRLQVPVEAVNMSLGLSLAATTPPQLLAGCPTVHQTCSNTYVVKGLCFPLGSLNRQOPQ 119
Db 80 VSLH-GSNHTSKYLGMTLATDAAKGSLACDGLSRTCDQNTYLSGLCYLFPQSLGPM 138
QY 120 KFEALRGCPQESDSTAFIVDVGSGSIIPHDFRAKEFISTVMEQLKKSKTLFSLMOYSEE 179
Db 139 QNRPAVQECMKGVLDVLFVFDGSQLDKRDKDFEKLFEPMKDMVKLSNTSYQFAAVQSTD 198
QY 180 FRHFTFKPE-QNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAPKILI 238
Db 199 CRTEFTFLDYVQKNKPDVLLGSQPMFLNTFRAINVYVAVHFVEESGARPDATKVLV 258
QY 239 LIITDGKFGDPLGYEDVIPLEADREG-----VIRYVIGVDGAFRSEKSRQELNIVASKP 291
Db 259 IITDG-----EASDKGNI SAADHITRIIIGIKGHFVSVQKTLHIHIFASEP 304
QY 292 PRDHVQINNPEALKTIQNLREKIPATEGTQTSSSSFEHMSQEGFSAITNGPILLS 351
Db 305 VBEFVKILDTFEKLDLTDLQRIYAIERTNRQDLTSFNMLSSSGISADLSKGHAVVG 364
QY 352 TVGSDWAGGVF-LYTSKEKSTFINMTRVDSMDMDAYLYGAAA-IILNRVQSLVLGAPR 409
Db 365 AVGAKDWAGGFLDLREDLQCATFVGQBPDTSDVRGGYLTGYTVAWMTSRSRPLLAAGAPR 424
QY 410 YOHIGLIVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHY 467
Db 425 YQHVGVLLFQAPAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGEAELLIGALPLFF 484
QY 468 EQTRGQVSVCPPLPRGORARWQCDVAVLYGEGQOPWGRFGAALTVLGDVNGDKLTDVAIGA 527
Db 485 GEORGRVFTY---QRRQSLFEMVSELODGPYPIGRFGEAITALTLDINGDRLTDVAVGA 541
QY 528 PQEEDNRGAVLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLT 587
Db 542 PLEE--QGANVIFNGKPG-GLSPQSPQRLEGTVGLSGIOWFGRSIHGVKDLGGDLADV 598
QY 588 VQAQGHVLLLRSQPVLRKALMEFNPREVARNVFECNDQVVKYKQAG-EVRVCLHVQKST 646
Db 599 VGAERGVVVLSSRPVDMVTILMSFPABIPVHEVECSYSAREEQKHGVLKACPRIKPLT 658
QY 647 RDLREGIQSVVTVLDALDSGRPHSRAVENETKNSITRQTVGLTGTCTELKLQLPNC 706
Db 659 PQ--FQGRLLANLTYTLQDGHMRMRSLGFLPDGSHLSGNTSITP-DKSCLDFFHFFPIC 715
QY 707 IEDVPSPIVLNPLNSL---VGTPLSFGN-LRPVLAEDAQRFLTALFFPEKNCNDNTCQ 762
Db 716 IQDLISPINVSLNFSLEEBGTPRDQKRAMQPIRLPSIHTV-TKEIPFEKNCKEDKKCE 774

Tue Nov 9 12:56:10 2004

763 DDLSITFSRMSDCLVVGSP-----REFNVTVTVNDGSDSYRTQVTFPPFLDLSRKV 816
 775 ANLTSSPAPS-----GFLRLMSSASLAVETLNSGDAYVWRLDLPFRGLSFRK 827
 817 STONORSQRWELACESASSTVSGAL-KSTSCSINHPIFPENSEVTFNITDVSAS 875
 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQVMNTLNS 882
 876 LGKLLKANVTSEN-NMPTNKTEFQELPVKYAVTVMTSHGVSPKYITNFASENTSR 934
 883 WEDFVELNGTVHCENENSLQEDNSAATHIPVLYPVNLIKQENSTLYISFTPKGPKTQ 942
 935 VMQHQYQVSNLQORSLPISLFLVPVRLNQTIVMDREQ---VTFSENLS-----TCHTK 986
 943 QVQHVYQV---RIQPSAYDHNMT-LEALGVPRPHSEDLITYTWSVQTDPLVTCHE 996
 987 E-RLPSSHDFLAELRKAPVNCSTAVCORIQCDIPFGIOEEFNATLKGNSLSPDWYIKTS 1045
 997 DLKRPSS---AEQPCLPGV-----QFPCPIV---RWEILLQVGTIVELSKEIKAS 1042
 1046 HNHLIVSTAEILFNDVSVFTLLPGQAFVRSQTKVEPEVFPNPLPLIVGSSVGGLLLL 1105
 1043 -STLSLSSLSVSFNSKHPHYLSKA-SEAQLVKVDLIHEKMLHYVVLGGIGLVLL 1100
 1106 ALITAAALYKLGFFKQYKXMM-SRGPPGAP 1136
 1101 FLIFALYKVGFFKRLKERMADGGVNGSP 1132

RESULT 6
 A;Species: Homo sapiens (man)
 C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
 C;Accession: A53213
 R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
 J. Biol. Chem. 269, 6016-6025, 1994
 A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. U
 A;Reference number: A53213; PMID:94164962; PMID:8119947
 A;Accession: A53213
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1179 <SHA>
 A;Cross-references: UNIPROT:P38570; GB:L25851; NID:G457244; PID:G457245
 C;Genetics:
 A;Gene: GDB:ITGAE
 A;Cross-references: GDB:330801
 A;Map position: 17p13
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1153; DB 2; Length 1179;
 Best Local Similarity 29.1%; Pred. No. 1.2e-71;
 Matches 342; Conservative 214; Mismatches 449; Indels 172; Gaps 39;

45 GSLYQCDYSTGS--CEPI-RLQVP-----VPAVMSLGLSLAATSPOLLACPTVHQ 95
 65 GPLHRCSLVQDEILCHPVEHVPKGRHGVTVVRSHHGVLICI-----QVLVRP--HS 117
 96 TCSNTYVKGCLFQSNLRQPO-----
 118 LSELT---GTCSLLGDLRLRQAQNFLENLLDPADVTDGDCYSNKEGGEDDVNTA 174
 120 KPEALRGCPQED-----SDIAFLVDGSGSIIPHDFFRAKEFIPTVMEQL-- 164
 175 RRRALKEEEDKEDEDEEEAGTEAIIILDGSSIDPPDFQAKDFISNMNRYE 234
 165 KSKTLFLSMQYSEEFRIHTFKFQNNPNRSLIKPTQLLQRTTATGIRKVVRELFN 224
 235 KCFECNFALVQYGVITQTEFLDRSDQVMSLARVQNTQGVSVTKTASAMQVLSIFT 294
 225 ITNGARKNAFKIILLITDGEKFGDPLGYEDVFEADREGVIRYVIGVGAFAFRSEKSRQEL 284

295 SSHGSRKSKVMVLTGIGIFEDPLNLTTVINSPMQGVVERFAIGVGEFFKSARTAREL 354
 285 NTVASPPRPHVFIQINNFEALKTIQNLKEKIFAIEGTQTGSSSSSEHEMSQGFSAAIT 344
 355 NLIASPPDETHAFKVTNNMALDGLSKLNIILISMEGT---VGDALHYQLAQIGFSQAIL 411
 345 SNGP-LLSTVGSYDWAAGGVFLY--TSKEKSTFFINMTVRVDSMDNDA---YLGYAAAILRN 398
 412 DERQVLGAVGAFDNGSGGALLYDTRSRGRFLNQTAATAAADAEEAAQSYLYGAVAVLHKT 471
 399 RVQSLVGLAPRYQHIGLIVAMFR-ONTGMWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFLVLEGEQMGYSFGSELCPVDIMDMSGT 529
 457 DLVLIGAPHYEQTRGGQVSVCPILQORARMQCDVLYGEGQGPWCFGAALTVLGDVN 516
 530 DFLVAAAPFVHVHGEGRVYVYRLSE--QDGSFLARILSHGPGFTNARPFGAMAAMDLS 588
 517 GDKLTDVAIGAP---GEEDNR--GAVLPHGTSGSGISPSHSQRIAGSKLSPRLQYFGQ 570
 589 QDKLTDVAIGAPLEGFAGDDGASFGSVIYNG-HWDGLSASPSQIRASTVAPGLQYFGM 647
 571 SLSGGQDLTWDGLVDLTVGAGQHVLLRSQPLRVKAIMPEPNPREVAENVECDNQVYKG 630
 648 SMAGGFDISGDLADITVGLTQAVVFRSPVVELKVSMAFTPSALP-----IGF 697
 631 KEAGEVAVLHVOKSTRDLREGQIQSVTVYDIALDGRPHSRVAVFNETKNSTRQTQVL 690
 698 NGVNVRLCFEI--SSVTASGSLREALNFTLVDVGKQRRLQCDVRSCLGLRWS 756
 691 GLTQTCETLKLQLEN---CIEDVSPVILKLNFLSVGTPLSAPFNGLRPLVLAEDAQLRFT 746
 757 SSGQLCEDL-LMTEGELCEEDCFNSASVKVSYQL-QTPBGTQDHPQPILDRYTEPPAI 814
 747 ALFPFCKNCGNDITCQDDLSTTFPMSLDCLVGGPREFTVTVVRNDGEDSYRQVTF 806
 815 FQLEVERACKNKLFCVAELQLA-TTVSQQLVWGLTKELTLNINLTNSGEDSYMTSMALN 873
 807 FPLDLSYRKVSTLQNSQSRWRWELACESASSTEVSGALKSTCSINHPFFENSEVTFNI 866
 874 YPRNLQ-----LKRMQKPPSPNIQCDPQPV---ASVLINCRIGHVPL-KRSSAHVSV 923
 867 TFDVDSKASLGNKLLKANVTSEN---MPRTNKTEFQ---LELPVKYAVTVMTSHGV 919
 924 VQLEENAFPNRTADITVTINSNERRSLANETHLQFRHGFVAVLSKPSIMYVNTQQGL 983
 920 S--TKYLNFTASNTSRVWQHQQVSNLQORSLSPLSLVLPVRLNQIVMDRPOVTFSE 977
 984 SHHKEFLFHVHGEN---LFGABYQ-----LQICVPTKLRGLQVAAVKKLTRTQ 1028
 978 NLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCORIQCDIPFGIOEEFNATLKGNSL 1037
 1029 ASTVCTWQEOACAYSS-VQHVFEWHSVCVTA-----SDXENVTVAAEIS 1073
 1038 FQWYIKTSHNHLIIVST-----ABILFNDVSVFTLLPGQAFVRSQTKVEPEVFPNPL 1091
 1074 WD-----HSELLKDVTELOILGEISFNKSLYEGNAENH--RTKITVVFLLDKYHSL 1125
 1092 PLIVGSSVGGLLLLALITAAALYKLGFFKQYKXMMSE 1128
 1126 PIIIKGSSVGLLIVLIVILEFKGFFKRYKQQLNLE 1162

RESULT 7
 A45226
 Integrin alpha-1 chain - human (fragment)
 C;Species: Homo sapiens (man)
 C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: A45226
 R;Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
 J. Biol. Chem. 268, 2989-2996, 1993
 A;Title: Expression of native and truncated forms of the human integrin alpha 1 subuni

A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1151 <BRI>

A:Cross-references: UNIPROT:P56199

A:Experimental source: hepatoblastoma cell line HepG2

A>Note: sequence extracted from NCBI backbone (NCBI:P124326)

F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

```

Query Match      18.8%; Score 1103.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 3.3e-68;
Matches 344; Conservative 212; Mismatches 486; Indels 195; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FGQSVVQL---QGRVVVVGAPQEIIVAANORGSLYQCDYSYGS 56
Db 1 FNVDKVSMTFSGPVEDMFGYTVQYQYENEEKWLIGSLVQGPKNRTGVDYKCPVGGRGE 60

QY 57 CEP-IRLOVPVEA-----VMSLGLSLAAATTSPPQLACGPTVHQTCSENTYVKG 106
Db 61 SLPCVKLDLPVNTSIPNVTEVENMTFGSTL-VTPNPGGFLACGPLYAYRCGLHYTGI 119

QY 107 CFLFGSNLRQOPKQPFPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLK- 165
Db 120 CSDVSPTFQVNVNSTAP--VQECSTQ-LDIVIVLDCGNSIYPWDSVTA--FLNDLLKMDI 174

QY 166 -KSKTFLSLMQYSEEFRIHFTFEKFNPNPRSLIKPITOLLGR-THTATGIRKVVRELP 223
Db 175 GPKQTQGVIGYQYGENVTHEFNLNKYSSTEEVLVAKKIVORGQOTWTALGTDTPARKEAF 234

QY 224 NITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGVGDADR-----SE 278
Db 235 TEARGARRGVKVMVITVDGESH-DNHRLLKVIQDCEDENIQFSAIILGSYNRGNLSTE 293

QY 279 KSRQELMTVASKPRDRHVFQINNFPEALKTIQONLREKIFAIEGTQTGSSSFHEMSQ 338
Db 294 KFVEEIKSIASEPTEKFFNVSDLAIVITKLTGERIFALEATADQSAASFEMESQ 353

QY 339 FSAAITSNGLPLTVSGSYDMAGGVFLYTSKE-----KSTF-INMTVDSMDNDAYLGAA 392
Db 354 FSAHSQDMWMLGAVGAYDNWGVVWQKASQIIPNTTNNVSTKNEPL-ASYLGTV 412

QY 393 AITLNRVQSL-VLGPARYQHIGLVAMPFRONTGWESNANVKGTQICAYGASLCSVDVD 451
Db 413 NSATASSGDVLYTAGQPRYHNTQVLIYRMEDGNIKILQLSGEQIGSYFSLITTDID 472

QY 452 SNGSTDVLVLGAPHY-----YEQTR-GGVSVCLPRGQARWQCDVAVLY 495
Db 473 KQNTDILLVGPMPYMGTEKEBGQYVYVVALNQTRFQYQMSLEPIKQTCSSRQHNSCTT 532

QY 496 GEQGPWG-RFGAALTVLGVDNGDKLTDVAIGAPGEDNKGAVYVLPFGTSGSGISPSHSQ 554
Db 533 ENKNEPCGARFGTAIAAVKDLNLDGENDIVIGAPLEDHGGAVIYTHG-SGKTIRKEYAQ 591

QY 555 RTAGSKLSPRLQYFGQSLSGGDLTMDGLVDLTVGAQGHVLLRSQPLVRKAIMFNPR 614
Db 592 RIPSQGDGKTLKFFGOSIHGEMDLNGDLTDTVIGLGAALFWSRDAVAVKVTMNFEN 651

QY 615 EVARNVFECDNQVVKGEAG--EVRVCLHVO-KSTDRLRBEOIQSVVYDIALDSGRPH 671
Db 652 KNIQKKNCH---MEGTEVCINATVCFEVKLKSKEDTIYADLQ---YRVLDSLRLQI 704

QY 672 SRAVFNET-----KNSTRQQTQLGLTQTCFTKLQLENCTEDPVPVILRLNLSVLT 725
Db 705 SRSFFSGTQERKQVORNIYRKSEC-----TKISFYLMDKHDFQDSVR-----ITLDFNLT-D 755

QY 726 PLSAFNLRLPVLAEADQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVWVGPRE- 784
Db 756 PENG-----PVLDDSLPNSVHEIYIPAKCKGKKEKICISDLSHVATTEKDLIVRSQNDK 810

QY 785 FNVTVTVRNDGEDSYRTQVTFPPDLDLSYRKVSTIQNQRQSRWRLACBSASTVSGAL 844
Db 811 FNVSLTVKMTKDSAYNTRTIVHYSPLNLFVSGIEAIQK-----SCBSN----- 853

```

```

QY 845 KSTSCSINHPIPPENSEVTFNITEDVDGSKASLGN-KULLKANVTSENMMRPTNKTEFOLE 903
Db 854 HNTCKGVGFLRRGEMVTFKILFQNTSYLMENVTIYLSATSDSEBPPETILSDNVNIS 913

QY 904 LPVKYAVVMVTSCHGVSTKYLNFPTASNTSRVMQHOYQVSN-----LGQR8-----L 950
Db 914 IPVKYEVGLQYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSPMPDEL 972

QY 951 PISLVF-----LVPRLNQTIVWDRQVTFSENLSSSTCTKE-----RLPS 991
Db 973 KLSISFPNMTSNGYVLYPTGLSS-----SENANCRPHIFEDPFINSKGMWT 1021

QY 992 HSDFLAELRKAPVNVCSIAVCQIQDPIPFQIOE-----EFNATLK 1033
Db 1022 STD--HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILMKPTFIKSYFSSINLTIR 1078

QY 1034 GNLSFDWIKTSHNHLIIVSTABILLFNDVSFTLLPGQAFVRSQTETKVEPFEVNPDLPL 1093
Db 1079 GEL-----RSENASLVLSN-----QKRELATQISKDGLGRVPL 1114

QY 1094 --IVGSSVGGILLIALITALYKLGFFPKROYKDMME 1128
Db 1115 WVLLSAFAGLLJLLMLLILALWKIGFKRPLKKMKEX 1151

RESULT 8
I45914
integrin alpha 2 subunit - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin
A:Reference number: A54402; MUID:94193647; PMID:7511592
A:Accession: I45914
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: UNIPROT:P53710; GB:L25986; NID:G439695; PIDN:AB59255.1; PID:G439696
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match      18.5%; Score 1085; DB 2; Length 1170;
Best Local Similarity 27.6%; Pred. No. 6.6e-67;
Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

QY 1 ENLDTENAMTFQ-ENARGGQSVVQL---QGRVVVVGAPQEIIVAANORGSLYQCD- 54
Db 19 YNVGLPKAKIFGSPSEQGYAVQQFINKGNMLLVGSPWSPFKRMGMDVYKCPVDLST 78

QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKG 107
Db 79 TTCEKLNLTSTSMNSVTEKTNMSLGLTLTRNVGTGGFLTCGLPWAQCGSQYTTGVC 138

QY 108 FLFGNLRQOPKQPFPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLK- 165
Db 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVWCDSENSIYPWD--AVKNELEKFXVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTFEKFNPNPRSLIKPITOLL-----GRTHATGIRKVVRE 221
Db 195 PTKTQMGTLQYANNPRVFNFLNFTKSKD---EMIKATQTFQYGGDLTNTFKAIOVARDT 251

QY 222 LFNITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRVIGV-----GDAFR 276
Db 252 AYSTAAGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKNILRFGIAYLVGLYLRNLD 310

QY 277 SEKSRQELMTVASKPRDRHVFQINNFPEALKTIQONLREKIFAIEGTQTGSSSFHEMSQ 336
Db 311 TKNLIKETAKIASIPTEHFFNVSDLAIVITKLTGERIFALEATADQSAASFEMESQ 369

QY 337 EGFSAAIT--SNGPLSLTVSGSYDMAGGVFLYTSKEKSTFPINNT--RVDSMN-DAYLGYA 391
Db 337 EGFSAAIT--SNGPLSLTVSGSYDMAGGVFLYTSKEKSTFPINNT--RVDSMN-DAYLGYA 391

```


Db 370 VFSAEYSPQNNILMLGAVGAYDWSGTVVQKTHGHLIFSKQAFEQILQDRNHSSYLIGS 429
Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRONTGMWENANV-----KGTQIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNVTVQSQRGQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDVLVIGAPHYEQTR--GGQVSVCLPRGORARWQCDVILYGBQGPWG 503
Db 485 CAVDVNDKTTDVLVAGAPMYNDLKKKEGRVYLFTIKG--ILNWH--QFLEGNGLUNA 541
Qy 504 RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYLPHGTSGSGISPSHSQRIAGS--KL 561
Db 542 RFGSAIALSDINWDGNDVIVGSPLENQSGAVIYINGHEG--IRLRSQKILGSDRAF 600
Qy 562 SPRLOYPQSLSGQDITMDGLVDLTVGAQGHVLLRSQPVLRVYKALMEFNPREVARNVF 621
Db 601 SSHLOYPGRSLDYGDLNGDSITDVSAGAFQGVVQLMSQSIADVSVDASFTEPKI--TL 658
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQISVVYDIALD---SGRPHSRAVEN 677
Db 659 NKNAEI-----KLKLCF-----SAKFRPTNNQNVAILVNTITDEQFSSRSVIRGLFK 707
Qy 678 ETKNSTRQTVLGTQFCE--TLKLQLPNCIEDPVSPVLRILNPSL--VGTPLSAFNL 733
Db 708 ENNERCLOXTMIVSQARCSHYIHIQEPS---DIISPLNLCMISLENPGT-----756
Qy 734 RPLVAEDAQRLLTALFPPEKNCNMNDNICODDLSTIF-----SFMSLDCLVGGPREFNVTV 789
Db 757 NPALAEYSETVKVFSIPPHKCGDDGVCISDLVNLVQQLPATQQQPFIVSNQNRKLTFSV 816
Qy 790 TVRNDGDSYQTVTFPPFLDLSYRKVSTLQNRQSRWRLACBSASST--EVSGALKSTS 848
Db 817 QLNKKESAYNTEIIVVDFSENLF-----ASWMPVDGTEVTCQIASQKSVT 864
Qy 849 CSINHIPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFQLELPVKY 908
Db 865 CNVGPALKSKQVTFITNFENLQ--NLQNASISFRALSQSEENMADNSVNLKLSLY 923
Qy 909 AVYVMVTSHGVSSTKYNLTASNTSRVMOHQVQVSNLQOR-----SLPISLVLV 958
Db 924 DAETHIT--RSTNINFEVSLDGNVSSV--HSPF--DIGPKFIFSVKVTGVSVPVMSA---976
Qy 959 PVRLNQTVWRPQVTFSEN--LSSTCHTKE-----RLPSHSDFLAE- 998
Db 977 -----SVLIHPIQYTKOKNPLMYLTGVHTDQAGDISCEAEINPLKIGOTS SVSFKSEN 1030
Qy 999 LRKAPVWNGSIAVCORIQDIPFGIQEENFATLKNLSFDWYIKTSHNHLIVSTABI- 1057
Db 1031 FRHIKELNCRATSCNIMCWLRLDQVKGBYFLNVSTRIWNGTFAASTFTVQLTAAAEID 1090
Qy 1058 LPNDSVFTL-----LPGGAFVRSQVETKVEPPE--VPNPLPLVGVSSVGGILLALITA 1110
Db 1091 TYNPQIVYIENVTIP-----LTIKPKHEKVEVPTGIIVGSVIAGILLALLALVA 1140
Qy 1111 ALYKLGFFKRYKQDM 1125
Db 1141 ILWKLGFKKRYEKM 1155

RESULT 9

S44142
VIA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
R:Accession: S44142
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VIA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PTDN:CAA82877.1; PID:g473098;
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;
Best Local Similarity 27.9%; Pred. No. 6.3e-66;
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;

Qy 1 ENLDTENAMTFQ--ENARGFGQSVVQL---QGSRRVVVGAPOEIVAAANQRGSLYQC--DYST 54
Db 27 YNVGLPKAKIFSFSSEHQFGYSVQQLTNPQGNLLVGSWSPGFPENRMGDIYKCPVLDPT 86
Qy 55 GSCPIRLQ-----VPVEAVNMISGLSLAATTSPPQLLACGTVHQTCSENTYVGLC 107
Db 87 ATCEKLNQNSASISNVTETKTNMSLGLTLTRNPGTGGFTLTCGFLMAHQCGNQYAYGIC 146
Qy 108 FLGSLNLRQPOQ---KPEPALRGCPQEDSDIAFLIVDGSGLIIPHDPRRAKEFISTVMQEQL 164
Db 147 ----SDVSPDFQLTSPSPAVQACPSL--VDVVVVCDESNIYP--WEAVKNFLVKFTGL 199
Qy 165 K--KSKTLFSLIMQYSEEFRIHTFKFQNNPNRSLIKPITQLLG-RHTATGTRKVVRE 221
Db 200 DIGPKKTQVALIQYANEPRIIFNLNDFETKEDMVQATSETROHGGDLTNTFRAIEFARDY 259
Qy 222 LFNITNGARKNAFKILITIDGKFGDPLGVEDVIPLEADREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSGRPCGATKVMVWVTDGESH-DGSKLKTVIQCCNDDEILRFGLAVLGYLNRNALD 318
Qy 277 SEKSRQELNVTASKEPRDRHVFIQINNFEALKTIQNLREKIFAIBGTQTGSSSPHEMSQ 336
Db 319 TKNLIKIKAIKATSTERYFPFNVADEAALKEKAGTLGQIIFSIETGVG--GDNQWENMAQ 377
Qy 337 EGPSA--AITSNGPLLLSTVSGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN--DAYLGYA 391
Db 378 VGFADYAPQNDILMLGAVGAFDMSGTIVQETSHKPVIFPKQAFQDVLQDRNHSFLGYS 437
Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRONTGMWENANV-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHFAVAGAPRANVTGQIVLYSVNK--QGNVTVIQSHRGQIGISYFSVLCSS 494
Qy 448 VDVDSNGSTDVLVIGAPHYEQTR--GGQVSVCLPRGORARWQCDVILYGBQGPWGFR 505
Db 495 VDVDDKTTDVLVAGAPMYNDLKKKEGRVYLFTIKGILNQHQ---FLEGEPTGNARF 551
Qy 506 GAALTVLGVNKGDKLTDVAIGAPGEEDNRGAYLPHGTSGSGISPSHSQRIAGSKLSR- 564
Db 552 GSALAAALSDINWDGNDVIVGSPLENQSGAVIYINGHGT--IRTKYSQKILGSGAFRR 610
Qy 565 --LQVPGQSLSGQDITMDGLVDLTVGAQGHVLLRSQPVLRVYKALMEFNPREVARNVPC 623
Db 611 HLQFFGRSLDYGDLNGDSITDVSIGALGQVILQMSQSIADVAIEALFTP-----660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDRREGQISVVYDIALD---SGRPHSRAVENET 679
Db 661 -DKITLLNKAADKITLKCFAEFRAFGQNNQV--AILFNMTLDADHSSRSVTSRGVREN 717
Qy 680 KNSTRQTVLGLTQTCET--LKLQLPNCIEDPVSPVLRILNFSLVGTPLSAFNLRPVL 737
Db 718 SERFLQKNVNVNEVQKCEHHHSIQKPS--DVVNPLDLRVDISLENPGTS-----PAL 768
Qy 738 AEDAQRLLTALPFPKNCNDNNICODDLISI-----TFSPMSLDCLVVGGPREFNVT 788
Db 769 EAYSETVKVFSIPFYKCGSDGICISDLILDVQQLPAIQTSF-----IVSNQNRKLTFS 823
Qy 789 VTVRNDGDSYQTVTFPPFLDLSYRKVSTLQNRQSRWRLACBSASST--EVSGALKST 847
Db 824 VILKRGESAYNVTVLAESENLF-----ASFSMPVDGTEVTCVSGSQSV 871
Qy 848 SCINHIPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTE--FQLELP 905
Db 872 TCDVGVYALKSEQVTFITNFENLQ--NLQNASINFAFSESQ--ETINKADNSVSLTIP 928
Qy 906 VKYAVYVMVTSHGVSSTKYNLTASNTSRVMOHQVQVSNLQOR-----SLPISLV 955


```

Db 929 LLYDAELHLT-RSTNINFEISSDENAPSVIK---SVEDIGPKPIFSLKVTAGAPVSM 984
Qy 956 FLV-----PVRNQTVIMDRPOVTP-SENLS 980
Db 985 LVTHIPIQYTKKPNLLYLTGIQDQAGDISCTAEINPLKLPHTA---PSVSFKNENFR 1040
Qy 981 STCHTKERLPSHDFLAELRKAPVWNCISIAVCORIQCDIPFGIQEENATLKNLSPDW 1040
Db 1041 ---HTKE-----LDCKRTTSCSNITCWLKDLHMKAEFIVNVTWVNR 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDSVTLIPGQAGFVRSQTETKVEPPEVNPPLPLVGSV 1100
Db 1081 FASFTQVQLTAARAEITHNPQLFVTEENAVTIPLIMKPTKAEVPT--GVIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAMTAGLWKLGFFKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence, revision 18-Sep-1992 #text_change 09-Jul-2004
A;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; UID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907
A;Note: The authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catmel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and
A;Reference number: A56793; UID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 10-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; UID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-25/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 26.9%; Pred. No. 1e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVQL---QGSRVVVGAPQEIETVAANORSLYQC--DYST 54
Db 30 YNVGLPEAKIFGSPSEQGYAVQVQFINPKGNLLVGVSPGPPENRMDGVTKCPVDLST 89
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGIC 107

```

```

Db 90 ATCEKLNLOTSTSPNVTMKTNMSLGLILTRNMTGGFLTCGLPWAQCCNQYTTGVC 149
Qy 108 FLFGSNLRQOPQKPEALRGCPQEDSDIAFLVDGSGSIIPHDERRAKEFIITVWEQLK-- 165
Db 150 SDISPDF-QLSASFSPATPCPSL-IDVVVCDSESNLYPMD--AVRNFLBKFVQGLDIG 205
Qy 166 KSKTLFSLMQYSEBFRHFTFKFQNNPNRSLIKPITQLLG-RTHATGATGIRKVVRELFN 224
Db 206 PTKTOVGLIOYANNPRVFNLTNYKTEEMIVATSTQSYGGDLTNTFGALQYARKVAYS 265
Qy 225 ITNGARKNAFKILLITDGEKFGDPLGYEDVIEADREGVIRYIGV-----GDAFSEK 279
Db 266 AASGRRSATKVVVVVTDGESH-DGSMKKAVIDQCNHNDILRFGIAVLGLNRLALDTKN 324
Qy 280 SROELNTVASKPRDRHVFOINNFEALKTIQNLRKIFAIEGTQTGSSSSFEHMSQEGF 339
Db 325 LIKEIKAIASIPTEFYFFNVSDAALLEKAGTIGEIQFISIEGTVOG-GDNFQEMSQVGF 383
Qy 340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKSKSTFINMT--RVSDSMN-DAYLGYAAAI 394
Db 384 SADYSSQNIDILMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSLYGVAA 443
Qy 395 ILNRNVQSLVLGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYFASLCSV 448
Db 444 ISTGESTHVFVAGAPRANYTGGIIVLYSVN-----ENGNITVIOAHRGDQIGSYFGSLCSV 498
Qy 449 DVDSNGSTDLAVLIGAPHYEQTR--GGQSVSCPLPRGORARWOCDAVLGYEQCPWGRFG 506
Db 499 DVDKDTITDVLVAGAPWYMSDLKKEGRVYVFIKKGILGQHO---FLEGEIGIENTRFG 555
Qy 507 AALTUVLDVNGDKLTVAIGAPDEENRGAVLYLFGHTSGSGISPHSQRISAGS--KLSPR 564
Db 556 SAIAALSDINMDGFNDVIVGSPLENQSGAVIYINGHGT-IRTKYSQKILGSDGAFRSH 614
Qy 565 LQYFGOSLGSQDLTWDGLVLTGAGHVLLLRQSPVLVRKAIMFNPREVARNVPECN 624
Db 615 LQYFGSLDGYGLNDGSDITDVSIGAFQVQVQLWSQSIADVAIEASTFTEKI--TLVNKN 672
Qy 625 DVVVGKEAGEVRVCLHVQKSTRDLRREGIQSVVTVYDLALD---SGRPHSAVENETK 680
Db 673 AQII-----LKLCF---SAKFRPTKQNNQVAIVNITLDDGFSRVTSGLFKENN 721
Qy 681 NSTRQTQVLGTQTC--ETLKLQLPNCIEDPVSPVLRNLNFSLVGTPLSAFGLRFLVA 738
Db 722 ERLCKQNMVYVNAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQRLFTALFPPEKNCNDNICQDDLSITF---SFMSLDCLVVGCPREFNVTVVRND 794
Db 773 AYSETAKVPSIFPHKDCGEDGJCISDLVLDVROIIPAAQOPFIVSNQKRLTFSVTLKNK 832
Qy 795 GEDSYRTQVTFPFLDLISYRKVSTLQNRQSRWSRLACESASST-EVSGALKSTCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASPSLPVDGTGTEVTCQVAASQKSVACDVGY 880
Qy 854 PIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKTQFQLELPVKYAVNV 913
Db 881 PALKREQQVTFITNDFNLQ-NLQQAQSLSFQALSESQEBENKADNLVNLKILPLYDAEI- 938
Qy 914 VTSHGVSVKYLNFTASENTSRVMOHQYQVSNLQOR-----SLPSLVLV----- 958
Db 939 ---HLTRNININFYEISSDGNVPSIVHSFEDVGPKFISLVKVTTSVPVSMATVHIHPQ 995
Qy 959 -----PVRNQTVIMDRPOVTP-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLKIGQT---SSVSFKSENER---HTKE- 1047
Qy 989 LPSHSDFLAELRKAPVWNCISIAVCORIQCDIPFGIQEENATLKNLSDFYIKTSHNH 1048
Db 1048 -----LNCRTASCSNVTCWLKDVHMKGEYFVNVVTRINWNGTFASSTFQT 1091
Qy 1049 LLIIVSTABI-LFNDVSFTLLPGQAFVRSQVETKVEPEVNP-----LP--LIVGSSVG 1100

```

1092 VQLTAAAEINTYNEIYVI-----EDNTVTIPLIMKPEKAEVPTGVIGSIIA 1141
1101 GLLLLALITAAALYKLGFFKQYKDM 1125
1142 GILLALLAVAILWKLGLFFKQYKDM 1166

RESULT 11
A35854
integrin alpha-1 chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C:Accession: A35854; S1243
R:gnatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
A:Reference number: A35854; MUID:90338125; PMID:2380249
A:Accession: A35854
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1180 <IGN>
A:Cross-references: UNIPROT:P18614; GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.1%; Score 1062; DB 2; Length 1180;
Best Local Similarity 27.4%; Pred. No. 2.7e-65;
Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

QY 1 FNLDENAMTFOENARG-FGQSVVQL---QGSRVVVGAPQEIIVAANORGSLYQCDYSTGS 56
DB 29 FNVVVKMSGSPGVEDMFGTVQYENEEKWVLGSLVPLVGQPKARTGVYKCPVGRER 88
QY 57 CEP-IRLQVPVEA-----VNMSGLSLAAATSPDQLACGPTVHQTQCEYNYVKG 106
DB 89 AMPCVKLDLPVNTSIPNVTETKENMTFGSTL-VTNPNGGFLACGLYAYRCGLHYTTGI 147
QY 107 CFLFGSNLRQOPQKPEALRCPQEDSDIAFLVDSGSIIPHDPERRAKEFTSTWEOQLK- 165
DB 148 CSDVSPTFQVNVSEAP--VQECSTQ-LDIVLDGSSNIYP--WESVTAFLNDLLKMDI 202
QY 166 -KSKTLFLMOYSEEFRIHFTKFPQNNPNRSLIKPITQLLG-RTHTATGIRKVVRELF 223
DB 203 GPKQTQVGIQYGENVTHEFNLNKYSSTEELVAANKIGRQGLQMTALGIDTARKEAF 262
QY 224 NITNGARKNAFKILITIDGKFGDPLGVEDVPEADREGVIRYVIGVDAPR-----SE 278
DB 263 TEARGARRGVKVMVITDGEH-DNYRLKQVIQDCEDENIQRFSAIILGHYNRGNLSTE 321
QY 279 KSRQELNTVASKPPRDHVPQINNFEALKTIONOLEKIFALEGTQTSSTSSFEHMSQEG 338
DB 322 KFVEEIKSTASEPTEKHFNFVSDEALVTIVKALGERIFALEATDQSAASFEMMSQTG 381
QY 339 FSAATITSGPLISTVGSVDWAGVFLYTSKEKSTFINMT--RVDSMDND---AYLGAYAA 393
DB 382 FSAHVSQDWMLGAVGAYDNQTVVMQANQMVIPHNTTFQTEPAKMWLEPLASYLGYTN 441
QY 394 IILNRVOSLVLGARYQHIGLVAMFRONTGMWESNANVKGTOIGAYGASLCSVDVDSN 453
DB 442 SATIPGDVLYIAGQPRYNHTGTQVWYIKMEDGNINTLQTLGGQIGSYFGSVLTTIDDKD 501
QY 454 GSTDLVLIGAPHY-----YEQTR-GGQVSVCPPLPRGQARWOCDAVLYGE 497
DB 502 SYTDLLLVGAPMYMGTEKEEQKVYVYVAVNQTRFYMQLSEPIROTCCSLKDNSTKEN 561
QY 498 CQQPWG-REGAALTIVLGDVNGDKLTDAITGAPCEEDNRGAVLYPHGTGSGTSPSHSQRI 556
DB 562 KNEPCGARGTAAIAAKVLDNVDFNDVIGAPLEDHAGAVYIYHG-SGKTIREAYAORI 620
QY 557 AGSKLSPLQYFGQSLSGGQDLTMCDGLVLTVGAQGHVLLRSQPVLRKATMEFNPRV 616
DB 621 PSGGDGKTLKFFGQSIGHGMDLNGDGLTDVTITGGLGGAALFWARDVAVVVKVTTNFPNKV 680

QY 617 ARNVFECNDQVVKGEAG--EVRVCLHVQ-KSTRDLRREGQIQSVVYDIALDSGRPHSR 673
DB 681 NIQKNCR---VEGKETVCINATMCFHVKLKSDEIYBADLQ-----YRVTLDLSRQISR 733
QY 674 AVNET-----KNSTRQTOVLGLTQTCETLKLQLPNCI-----EDPVSPIVLR 718
DB 734 SFESGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVRVTL 776
QY 719 NPSLVGTSLSAFNGLRPVLAEADAQRLFTALFPPEKNCNDNIQQDDLSITFSFMSLDCLV 778
DB 777 DFNLT-DPENG-----PVLDDALPNSVHEHIPAKCGNKERCISDLTLNVSTTERKSLLI 830
QY 779 VGGPRE-FNVTVTRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSRWRACESASS 837
DB 831 VKSQHDKFNVLTVKNGDSAYNTRTVVQHSNPLIFSGIEIQKD-----SCSEN-- 880
QY 838 TEVSGALKSTSCSINPIPFENSEVFNITFDVDSKASLGNKLL-LKANVTSENMPRTN 896
DB 881 -----QNTICRVGVPFLRAGETVTFKIIPOFTSHLSNAIHLSATSDSEEPLESIN 933
QY 897 KTEFQLELPVKYAV---YMWVTSHGVSST-----KYLNFTASENTSRVMQHOYQVSNL 945
DB 934 DNEWNTSIPKVEYGLQFYSSAEHLSVAANETIPEFINST--EDIGNEINVFYTKR 991
QY 946 GQRSLP---ISLVP-----LVPRLNQTVIND-----RP-----Q 972
DB 992 GHFMPPELQLSIGFPNLTADGYFVLYPIG-----WSSSDNVNCRPSLEDPPFGINSKK 1045
QY 973 VTFSS-----ENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDI--- 1019
DB 1046 MTISKSSEVLKRGTIQDCSSTC-----GVATITCSLLPSDLSQVNVSL 1088
QY 1020 ---PFFGIQEF---NATLKNLSFDWYIKTSHNHLIVSTAETILFNDSVETLLPGQAF 1073
DB 1089 LWKTFP-IRAHFSSNLNLTRELK-----SENSLTSSN----- 1123
QY 1074 VRSQTEKTVKFFVEFVNPPLPL--IVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128
DB 1124 RKRELAQISKDLGPRVPLWVILLAFAGLLMLLILALWKIGFFKRLPKKKMK 1180

RESULT 12
A41131
Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse
N:Alternate names: integrin alpha-4
C:Species: Mus musculus (house mouse)
C:Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004
C:Accession: A41131; S16742
R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzman, B.; Weisman, I.L.
J. Cell Biol. 115, 1149-1158, 1991
A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte
A:Reference number: A41131; MUID:92064645; PMID:1840602
A:Accession: A41131
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1039 <NEU>
A:Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:951484; PIDN:CAA37316.1; PID:951484
C:Superfamily: integrin alpha-4 chain
C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;
Best Local Similarity 22.7%; Pred. No. 1.4e-37;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDENAMTFO-ENARGFGQSVV-QLOGSR--VVVGAPQEIIVAAN---QRGSLYQCDY 52
DB 41 YNLDPENALLQGPSGTILFGYSVVLKSHGSKWLVGFTASWLSNAAVNVFAGVRCGI 100
QY 53 STG---SCEPIRLQVP-----VEAVNMSGLSLAAATSP-PQLIACG---PTVHQ 95
DB 101 RKNPNQTCQLQSGSGSGEPCGKTCLEERDNQWLGVTLRSQPGENGSIIVTCGRWKNIFY 160
QY 96 TCSNTYVKGCLFLFGSNLRQOPQKPEALRCPQEDSDIAFLVDSGSIIPHDPFRAXE 155

S06046

integrin alpha-4 chain precursor - human

N:Alternate names: CD49d; very late antigen-4 alpha chain; VIA-4 alpha chain

C:Species: Homo sapiens (man)

C:Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text_change 09-Jul-2004

C:Accession: S06046; A39355; D28018

R:Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.

EMBO J. 8, 1361-1368, 1989

A:Title: The primary structure of the alpha(4) subunit of VIA-4: homology to other integrins

A:Reference number: S06046; MUID:89356603; PMID:2788572

A:Accession: S06046

A:Molecule type: mRNA

A:Residues: 1-1038 <TAK>

A:CROSS-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1

R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.

Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991

A:Title: Characterization of the alpha-4 integrin gene promoter.

A:Reference number: A39355; MUID:91239513; PMID:2034655

A:Accession: A39355

A:Molecule type: DNA

A:Residues: 1-93 <ROS>

A:CROSS-references: GB:M62841

R:Takada, Y.; Strominger, J.L.; Hemler, M.E.

Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987

A:Title: The very late antigen family of heterodimers is part of a superfamily of molecules

A:Reference number: A94151; MUID:87204112; PMID:3033641

A:Accession: D28018

A:Molecule type: protein

A:Residues: 40-50, 'E', 52-53 <TA2>

C:Genetics:

A:Gene: GDB:ITGA4; CD49D

A:CROSS-references: GDB:128032; OMIM:192975

A:Map position: 2q31-2q32

C:Superfamily: integrin alpha-4 chain

C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmembrane

F:1-39/Domain: signal sequence #status predicted <SIG>

F:40-1038/Product: integrin VIA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.8%; Score 635; DB 2; Length 1038;

Best Local Similarity 25.3%; Pred. No. 1,2e-35;

Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;

QY 272 GDAFSEKSRQELNTVASKPPRD-----HVFQINFEALKT-----IQN 310

Db 121 GHTCLEERDNLGVTLRSQPGENGSIIVTCGHRWKNIFYIKENKLTPTGGCYGVPPDLRT 180

QY 311 QLREKI-----FALEGTQTCSSSFHEMSOEFSAAITNSGPLLSTVGSY 356

Db 181 ELSKIAPCYQDYVVKFENFA-----SCQAGISSFYTKDLIVMGAPGSS 225

QY 357 DWAGGVFLY---TSKEKSTFTINMTRVDSMDNDAYLGVA--AAAILRNVRVQSLVIGAPRYQ 411

Db 226 YWTGSLFVNIITNNKYAPLKDQNVKE---GSLYGVSVGAGHFRSQHTTEVVGAPQHE 282

QY 412 HIGLVAMFRONTGMWESNAVY----KGTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYY 467

Db 283 QIGKAYIF----SIDEKEENLILHEMKGKGLGASYFGASVCAVDLNDAGFSDL-LVGAPMQS 337

QY 468 EOTRGQGVSVCPPLRGQARWQC-DAVLYGQGGQPGWGRFGAALTVLGDVNGDKLTDVAIG 526

Db 338 TIREGRVFVY-INSGSAVMAMETNLVGSCKYA-ARFGESI VNLGDIDNDGFEDVAIG 395

QY 527 APGEDNRGAVYLFHGTSGSGISPSHSQRISAGSKLSPLRQYFGQSLSGQDLTDWGLVDL 586

Db 396 APQEDDLQAIYIYNGRA-DGISSTFESQRIEGLQISKLSMFGQISQIDADNNGYDV 454

QY 587 TVGA--QGHVLLRSQPLVRVKAIMEFNPREVARNPFCNDQVVKGKEAGERVVCLHVQK 644

Db 455 AVGAFRSDSALLTRPVVIVDASLS-HPESVNRKTFDC-----VENGWFSVCIDILT 506

Search completed: November 9, 2004, 12:19:34
Job time : 33.25 secs

QY	659	-----VYDIALDSG-RPHSRVAFNETKNSTRROTQVLGL--TOTCETKLQ	707
Db	498	CFTYGTNYLPDHIIDISYVTVDGSIANRRAMFVDNDMSEITRRLAVSTQFCDPLRAY	557
QY	703	LPNCIEDPVSPIVLRLNPSLNGTFLPSAFGN-----LRPVLAEADAORLFTALEPFFK	753
Db	558	VGNSIEDKLTPIKVTLODDL-----NNDESRLOPHEILPIDMATMSTQTKQVSION	609
QY	754	NGCNDNICODDLSTIFSPMSLDCLVGCGPREFNVTVTRNDGDSVXTQVTFEFLDLSY	813
Db	610	NCVN-NICIPOLDVTVT--PNLPNVIGTQBELIDVSLNRRGDAFQSSLSVVPYJGLQF	667
QY	814	RKVSTLQNRQSRGWRLACESASSTEVSGALKSTCSINHPFEPN-----SEVTFNIT	867
Db	668	VRL-----ERKANMDFSVTCSGSDS-----LRIITCDTGNPMVGKNILBFGTLSTFQVS	717
QY	868	FDVDS-----KASLGNKLLIKANVTSNNMPRNKTEFQLELPVKVAVYMWVTSHGVS	922
Db	718	GDKOSIEFYFAESSENS--EDPNLTENNELNMVTPVTDCTLLKSASYPEIWMYSTQSD	775
QY	923	YL-----NFTASENTSRYMHQYQV-----SNLGORSILPIS-----LVFLVPV	960
Db	776	YVVPFPKXNASEADIGMEVW-HUYEVNRNTGSSNAGEVSLNQWPKQKEDGBYLFYLLGI	834
QY	961	RLNQTVIWDREQ-----VTFSENLSTCTKRLPSHSDFLAELKAPVW	1005
Db	835	MTEEGVTQCLTQKANPEGVKLEPSTKAKLSNITTQVSGREKREPVAALAQTON--VI	892
QY	1006	NCSTAVCORIQCDIPFGIOEFFNAT-----LKGNSLFDWYIKTSHNHLIVSTAEILF	1055
Db	893	YCASDSCVLINCTI-----DEINASKSVVRILGRF--W-----ERTF	928
QY	1060	NDSVFTLLPGOGAFVRSQTEKVE--PFEPVNP-----IP-----	1095
Db	929	QKAVSELPVQAVTIAASAAAKTIPIYNTPLRDFSDSTKASTLVTTEELVPVPTIAW	988

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKROYKDMSEGGPPGABPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5846.5	99.5	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4470	76.1	1153	1	ITAM_MOUSE	P05555 mus musculu
3	4381	74.6	1151	2	Q9J130	Q9J130 rattus norv
4	3809.5	64.8	920	2	Q28984	Q28984 sus scrofa
5	3479	59.2	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3411	58.0	1162	1	ITAD_HUMAN	P20702 homo sapien
7	3329.5	56.7	1188	2	Q6KAS4	Q6KAS4 mus musculu
8	3329.5	56.7	1188	2	BAD21383	BAD21383 mus musculu
9	3320.5	56.5	1169	1	ITAX_MOUSE	Q9QXH4 mus musculu
10	3244.5	55.2	1161	1	ITAD_RAT	Q9QYF7 rattus norv
11	1583	26.9	1165	1	ITAL_BOVIN	P61625 bos taurus
12	1557.5	26.5	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1557	26.5	1166	2	Q6TYB8	Q6TYB8 bos taurus
14	1557	26.5	1166	2	AAQ90015	AAQ90015 bos taurus
15	1538.5	26.2	1163	1	ITAL_MOUSE	P24063 mus musculu
16	1534.5	26.1	1161	2	Q9RTV4	Q9RTV4 mus musculu
17	1524	25.9	1160	2	Q9R200	Q9R200 mus musculu
18	1409	24.0	1196	2	Q98TF1	Q98TF1 cyprinus ca
19	1359.5	23.1	1086	2	Q96HB1	Q96HB1 homo sapien
20	1358.5	23.1	1187	2	Q98TF0	Q98TF0 cyprinus ca
21	1378	21.7	927	2	Q8HZV0	Q8HZV0 bos taurus
22	1167.5	19.9	1167	2	Q88340	Q88340 rattus norv
23	1161.5	19.8	1167	1	ITAE_MOUSE	P38570 mus musculu
24	1153	19.6	1179	1	ITAE_HUMAN	P38570 mus musculu
25	1123	19.1	1167	2	Q88341	Q88341 rattus norv
26	1103.5	18.8	1151	1	ITAI_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1	ITAH_HUMAN	Q9UKX5 homo sapien
28	1085	18.5	1170	1	ITAI_BOVIN	P53710 bos taurus
29	1071	18.2	1178	1	ITAE_MOUSE	Q62469 mus musculu
30	1071	18.2	1178	2	Q6P1C7	Q6P1C7 mus musculu
31	1071	18.2	1178	2	AAH65139	AAH65139 mus musculu

RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MOI)			
DE	(Neutrophil adherence receptor).			
GN	Name=ITCAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

ALIGNMENTS

FT	SIGNAL	1	16	Integrin alpha-M.	
FT	CHAIN	17	1152		
Query Match 99.5%; Score 5846.5; DB 1; Length 1152;					
Best Local Similarity 99.2%; Pred. No. 0;					
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;					
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVQSGSLYQCDYSTGSCPEI	60		
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVQSGSLYQCDYSTGSCPEI	76		
QY	61	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGCLFGLSGSLRQOPQK	120		
DB	77	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGCLFGLSGSLRQOPQK	136		
QY	121	FPFALRGCPQEDSDIAPLVDSGSIIPHPFRRAKEPSTVMEQLKSKTFLPSLMQYSEEF	180		
DB	137	FPFALRGCPQEDSDIAPLVDSGSIIPHPFRRAKEPSTVMEQLKSKTFLPSLMQYSEEF	196		
QY	181	RIHFTFKFQNNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240		
DB	197	RIHFTFKFQNNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI	256		
QY	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	300		
DB	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	316		
QY	301	NFEALATIQNLREKIPALEGTQGTSSSPEHEMSQEGFSAATISNGPLLSVGSVDWAG	360		
DB	317	NFEALATIQNLREKIPALEGTQGTSSSPEHEMSQEGFSAATISNGPLLSVGSVDWAG	376		
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVGLAPRYQHILGLVAMFR	420		
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVGLAPRYQHILGLVAMFR	436		
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYQTRGGQVSCPL	480		
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYQTRGGQVSCPL	496		
QY	481	PRGORARWCCDAVLXGEOCPGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540		
DB	497	PRG-RARWCCDAVLXGEOCPGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	555		
QY	541	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLMGDLVDLTVCAGQHVLLRSQ	600		
DB	556	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLMGDLVDLTVCAGQHVLLRSQ	615		
QY	601	PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660		
DB	616	PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	675		
QY	661	YDLALDSGRPHRAVFNETHKSTRQTVGLTQTCETIKLQLPNCIEDPVSPVILRLNF	720		
DB	676	YDLALDSGRPHRAVFNETHKSTRQTVGLTQTCETIKLQLPNCIEDPVSPVILRLNF	735		
QY	721	SLVGTPLSAFGLNLRPLVAEDAQLRFTALPFPFKNCNGNDNICDDLSITPFSMSLCLVVG	780		
DB	736	SLVGTPLSAFGLNLRPLVAEDAQLRFTALPFPFKNCNGNDNICDDLSITPFSMSLCLVVG	795		
QY	781	GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASTEV	840		
DB	796	GPREFNVTVVRNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASTEV	855		
QY	841	SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNKTEF	900		
DB	856	SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNKTEF	915		
QY	901	QLELPKVAVYVMVTVSHGVSTKYLNFATASENTSRVWQHOYQVSNLQORSPLSLVFLVPV	960		
DB	916	QLELPKVAVYVMVTVSHGVSTKYLNFATASENTSRVWQHOYQVSNLQORSPLSLVFLVPV	975		
QY	961	RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRIQCIP	1020		

Db	976	RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNCISIAVCQRIQCIP	1035
QY	1021	FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI	1080
Db	1036	FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI	1095
QY	1081	KVEPFEVNPPLPLIVGSSVGGMLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1137
Db	1096	KVEPFEVNPPLPLIVGSSVGGMLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1152

RESULT 2

ID	ITAM_MOUSE	STANDARD;	PRT;	1153 AA.
AC	P05555; Q8CA73;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI).			
GN	Name=itgam;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=88312584; PubMed=3044779;			
RA	Pytela R.;			
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor.";			
RL	EMBO J. 7:1371-1378(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	STRAIN=CS7BL/6J; TISSUE=Spinal cord;			
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;			
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,			
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,			
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,			
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,			
RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,			
RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,			
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,			
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,			
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,			
RA	Kanai A., Kawaji H., Kasawara Y., Kedzierski R.M., King B.L.,			
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,			
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,			
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,			
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,			
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,			
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,			
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,			
RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,			
RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,			
RA	Tuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,			
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,			
RA	Shiaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,			
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,			
RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,			
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,			
RA	Birney E., Hayashizaki Y.;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs.";			
RL	Nature 420:563-573 (2002).			
RN	[3]			
RP	SEQUENCE OF 11-45 FROM N.A.			
RC	STRAIN=BALE/c; TISSUE=Spleen;			
RX	MEDLINE=86287312; PubMed=2942940;			
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,			
RA	Larson R.S., Roberts T.M., Springer T.A.;			
RT	"A partial genomic DNA clone for the alpha subunit of the mouse			

complement receptor type 3 and cellular adhesion molecule Mac-1-";
 Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).
 [4]
 SEQUENCE OF 17-28.
 MEDLINE=85188276; PubMed=3887182;
 Springer T.A., Teplov D.B., Dreyer W.J.;
 "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon.";
 Nature 314:540-542(1985).
 -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
 adhesive interactions of monocytes, macrophages and granulocytes
 as well as in mediating the uptake of complement-coated particles.
 It is identical with CR-3, the receptor for the iC3b fragment of
 the third complement component. It probably recognizes the R-G-D
 peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
 fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
 of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
 mast cell development and in immune complex-mediated
 glomerulonephritis. Mice expressing a null mutation of the alpha-M
 subunit gene demonstrate increase in neutrophil accumulation, in
 response to a impaired degranulation and phagocytosis, events that
 apparently accelerate apoptosis in neutrophils. These mice develop
 obesity.
 -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
 associates with beta-2.
 -!- SUBCELLULAR LOCATION: Type I membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1;
 IsoId=P05555-1; Sequence=Displayed;
 Name=2;
 IsoId=P05555-2; Sequence=VSP_010473;
 Note=No experimental confirmation available;
 TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 granulocytes.
 -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 with I-domains do not undergo protease cleavage.
 -!- SIMILARITY: Belongs to the integrin alpha chain family.
 -!- SIMILARITY: Contains 7 FG-GAP repeats.
 -!- SIMILARITY: Contains 1 VWFA domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; X07640; CAA30479.1; -;
 EMBL; AK039444; BAC30350.1; -;
 EMBL; M14293; AAA39484.1; -;
 PIR; S00551; S00551.
 HSSP; P11215; 1BHQ.
 MGI; MGI:96607; Itgam.
 GO; GO:0009897; C:external side of plasma membrane; IDA.
 GO; GO:0007155; P:cell adhesion; IMP.
 GO; GO:0045123; P:cellular extravasation; IMP.
 GO; GO:0030593; P:neutrophil chemotaxis; IMP.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR02035; VWF_A.
 Pfam; PF01839; FG-GAP; 3-
 Pfam; PF00357; Integrin_alpha; 1.
 Pfam; PF00092; VWA; 1.
 PRINTS; PR01185; INTEGRINA.
 PRINTS; PR00453; VWFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS02034; VWA; 1.
 Alternative splicing; Calcium; Cell adhesion;
 Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
 Signal; Transmembrane.

FT	SIGNAL	1	16	Integrin alpha-M.
FT	CHAIN	17	1153	Extracellular (Potential).
FT	DOMAIN	17	1105	Potential.
FT	TRANSNEM	1106	1129	Cytoplasmic (Potential).
FT	DOMAIN	1130	1153	FG-GAP 1.
FT	REPEAT	31	84	FG-GAP 2.
FT	REPEAT	85	163	VWFA.
FT	DOMAIN	164	350	FG-GAP 3.
FT	REPEAT	337	400	FG-GAP 4.
FT	REPEAT	401	452	FG-GAP 5.
FT	REPEAT	454	515	FG-GAP 6.
FT	REPEAT	517	575	FG-GAP 7.
FT	REPEAT	580	632	Potential.
FT	CA_BIND	465	473	Potential.
FT	CA_BIND	529	537	Potential.
FT	CA_BIND	592	600	GPPFR motif.
FT	SITE	1132	1136	By similarity.
FT	DISULFID	66	73	By similarity.
FT	DISULFID	105	123	By similarity.
FT	DISULFID	654	711	By similarity.
FT	DISULFID	770	776	By similarity.
FT	DISULFID	999	1023	By similarity.
FT	DISULFID	1028	1033	By similarity.
FT	CARBOHYD	58	58	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	86	86	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	391	391	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	734	734	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	772	772	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	801	801	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	881	881	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	907	907	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	941	941	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	980	980	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	994	994	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1022	1022	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1045	1045	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1051	1051	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1076	1076	N-linked (GlcNAc. .) (Potential).
FT	VARSPIC	453	569	Missing (in isoform 2). /Ftid=VSP_010473.
FT	CONFLICT	37	37	N -> S (in Ref. 2).
FT	CONFLICT	683	683	V -> G (in Ref. 2).
FT	SEQUENCE	1153	AA; 127480	MW; 178DB988AECB0343
QY	Query Match	76.18	Score 4470; DB 1; Length 1153;	
Db	Best Local Similarity	73.94	Pred. No. 7.9e-287; Indels 2; Gaps 2;	
Db	Matches 841; Conservative 146; Mismatches 149;			
QY	1	FNLDTENAMTFQENARGFGQSVVQLGSRVVGAPQIVAAQNRGSLYQCDYTGSCPEI	60	
Db	17	FNLDTEHPMTFQENAKGFGQNVVLGGTSVVVAAPQAKAVNQTGALYQCDYSTSRCHPI	76	
QY	61	RLQVPVAVNMUGLSLAATTSPPOLLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK	120	
Db	77	PLQVPPEAVNMUGLSLAVSTVPQOLLACGPTVHTQNKENTYVNGLCYLFGLNLRPPQQ	136	
QY	121	FPEALRCQPEDSDIAFLVDGSGSIIIPDFERRAKEFTSTVMEQLKSKTLFSLMOYSEEF	180	
Db	137	FPEALRCQPEQSDIVFLIDGSGSINNIDFQMKKEFTSTVMEQFKSKTLFSLMOYSEEF	196	
QY	181	RIHFTTFKFNPNPRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240	
Db	197	RIHFTFNDFKPNPSRSHVSPIKQLNGRTKTASGIRKVVRELFNITNGARKNAFKILVVI	256	
QY	241	TDGEKFGDPLGYEDVIEADREGVIRVIVGVGDFAFRSEKSPQELNTVASKPRDHVFOIN	300	
Db	257	TDGEKFGDPLDYKDVIEADRAGVIRVIVGVGNAFNKPQSRRELDITASKPAGHVFQVD	316	
QY	301	NFEALKTIONCLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG	360	
Db	317	NFEALNTIONLOEKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG	376	

QY 361 GVFLYTSKEKSTFTNMTVRDSDMDNDAYLGVAIAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
 Db 377 GAFLYTSKDKVTFINTTRVSDMDNDAYLGVAIAAILRNVRQSLVLGAPRYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKTQTGAIFGASCLSDVDVDSNGSTDLVLGAPHYETQTRGGQVSCVPL 480
 Db 437 ENFTWEPHTSISKQSGSYFGASCLSDVDMDADGNTNLILGAPHYETKTRGGQVSCVPL 496
 QY 481 PRGORARWQCDVLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
 Db 497 PRG-RARWQCEALHGDQHPGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYIF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRQ 600
 Db 556 YGASIASLSASHSHRIIAGHFSGLQYFGQSLSGGKDLTMDGLMDLAVGAGQHLLLRQA 615
 QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDRLRREGIOQSVVT 660
 Db 616 PVLRLIATMBFSPKKVARSVFACQEVLRKNDAGEVRVCLRVKNTDRLRREGIOQSVVT 675
 QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNF 720
 Db 676 YDLALDPGRSVVRAPFDETKNGILRRIRVFGTQKCTETLKLQLPDCVDSVSPILRLNY 735
 QY 721 SLVGTPLSAFNLRLPVLAEADAORLTALFPFEKNCNDNICQDDLSITFSFMSLDCILVVG 780
 Db 736 TLVGEPLRSFNLRLPVLADQAFRTAMFPFEKNCNDNICQDDLSITMSAMGLDILVVG 795
 QY 781 GPREENVTVRDGEDSYTQTFPPDLSTVRKYSTLQNRQSRWSRL-ACESASSTE 839
 Db 796 GPQDFNMSVTLRNDGEDSYTQTVVYPSGLSVRKDSASQNPULTKKPWFVKPAESSSSE 855
 QY 840 VSGALSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSSNNPRTNKT 899
 Db 856 HGCAKSTWNIHNPFPANSEVTFNITFDVDSHAFGNKLLKAI VASENNMSRTHKT 915
 QY 900 FOLELPVKVAYMVVTSHGVSSTKYLNTASNTSRVQHQYQVSNLQSRSLPISLVFLVP 959
 Db 916 FOLELPVKVAYMIVTSDSSIRYLNFTASEMTSKVIHQYQFNLLQSRSLPVSVVFWIP 975
 QY 960 VRLNQIVWDRPQVTSNLSSTCHTKERLPSHSDFLAEIRKAPVNVCSIAVQRIQCDI 1019
 Db 976 VQNNVTVDHPQVIFSONLSSACHTEQSPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035
 QY 1020 PFGIOEEFNALKGNLSPDWLYKTSHNLLIYSTAEILFNDSVFTLLPQCGAFVRSQTE 1079
 Db 1036 PSNTQEIFNVILKGNLSPDWLYKTSHGHLLSSTAEILFNDSAFALLPQGESVRSKTE 1095
 QY 1080 TKVEPEVNPPLIYGVSSVGGLLALITALIYKLGFFKQYKDMWSEGGPPGAPQ 1137
 Db 1096 TKVEPEVNPPLIYGVSSIGGLVLLALITAGLYKLGFFKQYKDMWNEAAPQDAPQ 1153

RESULT 3
 Q9J130
 ID Q9J130 PRELIMINARY; PRT; 1151 AA.
 AC Q9J130;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Integrin beta 2 alpha subunit.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fathallah D.M. Sr., Zerria K. Jr.,
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 DR EMBL; AF268593; AAF81280.1; -.
 DR HSSP; P11215; 1BHQ.

DR GO: 0008305; C:integrin complex; IEA.
 DR GO: 0007160; P:cell-matrix adhesion; IEA.
 DR GO: 0007229; P:integrin-mediated signaling pathway; IEA.
 DR InterPro: IPR000413; Integrin_alpha.
 DR InterPro: IPR002035; VWF A.
 DR Pfam: PF01839; FG-GAP; 2.
 DR Pfam: PF00357; Integrin_alpha; 1.
 DR Pfam: PF00032; VWA; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 DR Cell adhesion; Integrin; Transmembrane.
 KW

SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 2; Length 1151;
 Best Local Similarity 72.7%; Pred. No. 6.2e-281;
 Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 60
 Db 17 FNLDTENPMTFQENASGFGQSVIQLGETRVVVAAPQEVKAVNQTGALYQCDYSTNCRDPI 76
 QY 61 RLQVPEAVNMSLGLSLAATTSPQALLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
 Db 77 PLQVPEAVNMSLGLSLAATTVPQALLACGPTVHQNCKENTYVNGLCYLFSGNLLRQPOQ 136
 QY 121 FPEALRGCEQEDSDIAPLVDSGSIIPHFRRAKEFISTVMEOLKSKTKTFSIMQYSEEF 180
 Db 137 FPEALRGCEQESNIAFLIDGSSINTIDFQKKEFVSTVMDQFQSKTLFSLMQYSDEF 196
 QY 181 RIHFTFKFQNNPNPSLKIPTQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
 Db 197 RTHFTFENDPKRNPDPKSHVRPIQLNGRTKTASGIRKVVRELQKINGARDNAKILVVI 256
 QY 241 TQGEKEDPLGYEDNITPEADREGVIRYIVGVGDADFSEKSRQELNVTASKPPRDHVFQIN 300
 Db 257 TQGEKEDPLGYEDNITPEADREGVIRYIVGVGDADFSEKSRQELNVTASKPPRDHVFQIN 316
 QY 301 NPEALKTIONLRKEKIFALEGTQSGSSFEHMSQEGFSAAITSGNPLISTVGSVDWAG 360
 Db 317 NPEALNTIIRQKEKIFALEGTQSGSSFEHMSQEGFSAAITSGNPLISTVGSVDWAG 376
 QY 361 GVPLYTSKEKSTFNNTRVDSMDNDAYLGVAIAAILRNVRQSLVLGAPRYQHIGLVAMFR 420
 Db 377 GAFLYPSKDKASFINTTRIDSDMDNDAYLGVAIAAILRNVRQSLVLGAPRYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKTQTGAIFGASCLSDVDVDSNGSTDLVLGAPHYETQTRGGQVSCVPL 480
 Db 437 QNTGMWESNANVKTQTGAIFGASCLSDVDVDSNGSTDLVLGAPHYETQTRGGQVSCVPL 496
 QY 481 PRGORARWQCDVLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
 Db 497 PRG-RARWQCEALHGDQHPGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYIF 555
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLLRQ 600
 Db 556 HGASVASISTPHSQRIAGARFSPGLQYFGQSLSGGKDLTMDGLMDLAVGAGQHLLLRQA 615
 QY 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDRLRREGIOQSVVT 660
 Db 616 PVLRLIATMBFSPKKVARSVFACQEVLRKNDAGEVRVCLRVKNTDRLRREGIOQSVVT 675
 QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNF 720
 Db 676 YDLALDPGRSVVRAPFDETKNGILRRIRVFGTQKCTETLKLQLPDCVDSVSPILRLNY 735
 QY 721 SLVGTPLSAFNLRLPVLAEADAORLTALFPFEKNCNDNICQDDLSITFSFMSLDCILVVG 780
 Db 736 TLVGEPLRSFNLRLPVLADQAFRTAMFPFEKNCNDNICQDDLSITMSAMGLDILVVG 795

```
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLONORSORSRLACESASSTEV 840
Db 796 DSRDFDVSFVLRNDGEDSYTKVTCYYPGSLSVRKVSASONQFSKPPWRVIAE-PPSSSEG 854
QY 841 SGALKSTCSINHPPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 855 QGVLSKTIWINDHPPIFANSEVFNITFDVDSASLGNKLLKANVTSENNVSRDTKTES 914
QY 901 QLELPVKYAVYVWVTSHTGVTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 960
Db 915 QLELPVKYAVYVWVTSHTGVTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 974
QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELKAPVWVNCISIAVCORIQCIP 1020
Db 975 QINKVTIWDPRQVTFSENLSTCHTKERLPSHSDFLAELKAPVWVNCISIAVCORIQCIP 1034
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1080
Db 1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLIVSTAEILFNDVSFAFALLPGQETFKVKAQTEI 1094
QY 1081 KVBPFFVNPDLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1095 KVBPYTVHNPVLTIVGSSVGLLLALITAGLYKLGFFKQYKDMNEAGQDGPQP 1151
```

RESULT 4

```
Q28984 PRELIMINARY; PRT; 920 AA.
ID Q28984
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -2- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3-.
DR Pfam; PF00092; VWA; 1-.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS0234; VWFA; 1.
KW Cell adhesion; integrin; Transmembrane.
FT NON_TER 1
FT TER 920
FT SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.8%; Score 3809.5; DB 2; Length 920;
Best Local Similarity 78.9%; Pred. No. 3.2e-243;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;
QY 118 PQKPFALRGCPQEDSIAFLVDGSGSIIPHDFFRAKEFISTVMBQLKSKTLFSLMOYS 177
Db 1 PQKPFALRGCPQEDSIAFLVDGSGSIIPHDFFRAKEFISTVMBQLKSKTLFSLMOYS 60
QY 178 EEFRIHTFKEFQNNPNPSLIKPIQLLGRTHATGIRKVVRELFNITNGARKNAFKIL 237
Db 61 EDFYHTFTFNDKRNPSKLLVRPIQLLGRTHATGIRKVVRELFNITNGARKNAFKIL 120
```

```
QY 238 ILITDGEKFGDPLGYEDVIEPADREGVIRVIVIGVDAFRSEKGRQELNTVASKPPRRDHVF 297
Db 121 VVITDGEKFGDPLGYEDVIEPADREGVIRVIVIGVDAFNWKSRRELNTIASKPCGDHVF 180
QY 298 QINNFEALKTIONQLEKIPAIETGOTGSSSSSEHEMSQBFSAATNSGPLLSTVGSYD 357
Db 181 QVNNFEAVKTIQOLQEKTFATGOTGSSSSSEHEMSQBFSAATNSGPLLSTVGSYD 240
QY 358 WAGGVFLYTSKEKSTFINMTRVDSMDNDVLYGAAAILLRNVQSLVLAGAPRYQHIGLVA 417
Db 241 WAGGAFULHMPKDRVIFINTTRVDSMDNDVLYGAAAILLRNVQSLVLAGAPRYQHIGLVA 300
QY 418 MFRQNTGMBSNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQVS 477
Db 301 MFKQNSGAEKNADIKGSIQSYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQVS 360
QY 478 CPLPRGORARWQCDVLYGQGOQFWGRFGAALTIVLGVNVDGKLTVDVAIGAPBEDNRGAV 537
Db 361 CPLPQG-RAKWQCRVILCGEQHPWSEFGAALTALGDVWNGDKLTVDVAIGAPBEDNRGAV 419
QY 538 YLFHGTSGSIGSPSHSRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTWGAQHVL 597
Db 420 YLFHGTSELGSPSHSRIAGSKLSPRLQYFGOSLGGQDLTMDGLMDLAVGAQHVL 479
QY 598 RSQVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
Db 480 RSQVLRVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 539
QY 658 VVTVYDLALDSGRPHSRVAFNETKNSRRQTVLGLTOTCETLKLQLPNCIEDVPVPIVLR 717
Db 540 IITYDLALDPGRPHSRVAFNETKNSRRQTVLGLTOTCETLKLQLPNCIEDVPVPIVLR 599
QY 718 LNFSLVGTPLSAGNLPVLAEDAQRLLFTALFPFCKNGNDNICODDLSITFSFMSLDCL 777
Db 600 LNFSLVGTPLSAGNLPVLAEDAQRLLFTALFPFCKNGNDNICODDLSITFSFMSLDCL 659
QY 778 VVGPRFNFVTVVRNDGEDSYRTQVTFPPDLISYRKVSTLONORSORSRLACESAS 837
Db 660 VVGPRDLKVTLTVRNQGDSYRTQVTFPPDLISYRKVSTLONORSORSRLACESAS 719
QY 838 TEVSGALKSTCSINHPPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 897
Db 720 TEESTALKSTCSINHPPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 779
QY 898 TEFQLELPVKYAVYVWVTSHTGVTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFL 957
Db 780 TEFQLELPVKYAVYVWVTSHTGVTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFL 839
QY 958 VPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVWVNCISIAVCORIQC 1017
Db 840 VPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVWVNCISIAVCORIQC 899
QY 1018 DIPFFGQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1080
Db 900 DIPFFGQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFAFALLPGQETFKVKAQTEI 920
```

RESULT 5

```
ITAX HUMAN STANDARD; PRT; 1163 AA.
ID ITAX_HUMAN
AC P20702; Q8IVA6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Lue M5).
GN Name=ITCAX; Synonyms=CD11c;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=88166645; PubMed=3327687;
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
 RT "CDNA cloning and complete primary structure of the alpha subunit of a
 RL leukocyte adhesion glycoprotein, p150,95.";
 RN ENBO J. 6:4023-4028 (1987).
 [2]
 RX SEQUENCE FROM N.A.
 RP MEDLINE=90153906; PubMed=2303426;
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
 RL molecule.";
 RN J. Biol. Chem. 265:2782-2788 (1990).
 [3]
 RX ERRATUM.
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
 RL J. Biol. Chem. 265:12750-12751 (1990).
 [4]
 RX SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Tohiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [5]
 RX SEQUENCE OF 20-43.
 RX MEDLINE=87167596; PubMed=3549901;
 RA Miller L.J., Wiebe M., Springer T.A.;
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1
 RL and p150,95 leukocyte adhesion proteins.";
 RL J. Immunol. 138:2381-2383 (1987).
 CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 CC granulocytes.
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M81695; AAA51620.1; -.
 DR
 DR EMBL; M29165; -. NOT ANNOTATED CDS.
 DR EMBL; M29487; AAA51620.1; AUT_SEQ.
 DR EMBL; M29482; AAA51620.1; JOINED.
 DR EMBL; M29483; AAA51620.1; JOINED.
 DR EMBL; M29484; AAA51620.1; JOINED.
 DR EMBL; M29485; AAA51620.1; JOINED.
 DR EMBL; M29486; AAA51620.1; JOINED.
 DR EMBL; BC038237; AAH38237.1; -.
 DR PIR; A36584; RWHUIC.
 DR PDB; 1N3Y; X-ray; A=141-338.
 DR Genew; HGNC:6152; ITGAX.
 DR MIM; 151510; -.
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0004872; F:receptor activity; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0009887; P:organogenesis; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWFA.
 DR Pfam; PF01839; FG-GAP_3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWFA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWFA_1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS0234; VWFA; 1.
 KW 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 19
 FT CHAIN 20 1163 Integrin alpha-X.
 FT DOMAIN 20 1107 Extracellular (Potential).
 FT TRANSMEM 1108 1128 Potential.
 FT DOMAIN 1129 1163 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 ? FG-GAP 2.
 FT DOMAIN 165 351 VWFA.
 FT REPEAT ? 401 FG-GAP 3.
 FT REPEAT 402 453 FG-GAP 4.
 FT REPEAT 455 517 FG-GAP 5.
 FT REPEAT 518 576 FG-GAP 6.
 FT REPEAT 581 633 FG-GAP 7.
 FT CA_BIND 466 474 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 593 601 Potential.
 FT SITE 1131 1135 GFFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 655 712 By similarity.
 FT DISULFID 771 777 By similarity.
 FT DISULFID 848 863 By similarity.
 FT DISULFID 998 1022 By similarity.
 FT DISULFID 1027 1032 By similarity.
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).
 FT VARIANT 48 48 W -> R (in dbSNP:11574633).
 FT /FTID=VAR_018672.
 FT CONFLICT 209 209 T -> S (in Ref. 4).
 FT CONFLICT 251 251 T -> A (in Ref. 4).
 FT CONFLICT 469 469 T -> S (in Ref. 4).
 FT CONFLICT 490 490 E -> A (in Ref. 2).
 FT CONFLICT 547 547 G -> K (in Ref. 4).
 FT CONFLICT 756 756 D -> L (in Ref. 1).
 FT CONFLICT 819 819 I -> V (in Ref. 4).
 FT CONFLICT 1161 1163 SEK -> TPHYPQDNV (in Ref. 4).
 FT STRAND 150 150

```
FT TURN 160 161
FT HELIX 164 178
FT TURN 179 180
FT TURN 183 185
FT TURN 186 193
FT STRAND 197 201
FT STRAND 203 208
FT HELIX 212 216
FT TURN 217 218
FT STRAND 226 226
FT STRAND 228 236
FT TURN 237 240
FT HELIX 242 244
FT TURN 245 245
FT TURN 248 249
FT STRAND 251 258
FT STRAND 263 263
FT STRAND 269 278
FT TURN 279 280
FT STRAND 282 288
FT HELIX 290 293
FT TURN 296 297
FT HELIX 298 304
FT HELIX 310 312
FT STRAND 313 316
FT HELIX 319 325
FT HELIX 326 334
FT TURN 335 335
SQ SEQUENCE 1163 AA; 127886 MW; 836588A13B5C5DE8F CRC64;

Query Match 59.2%; Score 3479; DB 1; Length 1163;
Best Local Similarity 61.1%; Pred. No. 3.6e-221;
Matches 690; Conservative 141; Mismatches 292; Indels 6; Gaps 4;

QY 1 ENLDTENAMTFOENARGFGSVQVLOGSRVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 60
DB 20 ENLDTEELTAFRVDSAGFGDSVVQYANSVVVVGAPQIKITAAQNTGGLYQCGYSTGACEPI 79

QY 61 RLQVPEAVNMGLSLAATSPOLLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK 120
DB 80 GLQVPEAVNMGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCLFLGPT--QLTOR 137

QY 121 FPEALRGCPQSDIAFIVDGGSGIIPHDPRAKEFTSTVMEOLKXKTLFSLMOYSEEF 180
DB 138 LPVSRQCPQEQDIFVLIDGGSSISSRNFAFMNFRAVISQQRSTQPSLMQFNKP 197

QY 181 RIHFTFKFQNNPNSRIKPIITQLGRTHATGIRKVVRELFNITGARKNAFKILILI 240
DB 198 QHTFTFEFRRTSNFLSLASVHLQGGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

QY 241 TDGKFGDPLGYEDVIPPADREGVIRVIGVGDAPRSEKSOELNTVASKPRDHRVQIN 300
DB 258 TDGKKEGSDLYDKVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKE 317

QY 301 NFEALKTIQNLREKIPIAIEGTQTGSSSPHEMSQEGFSAATISNGPLLTSTVGSYDWAG 360
DB 318 DFDALDKIQNLQKKEKPIAIEGTETSSSFELEWAQEGFSAVTPDGPVLCAVSFTWSG 377

QY 361 GVFLYTSKEKSTFINMTVRVSDMNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
DB 378 GAFLYPENMSPTFINMSQENVDMDSYLGYSYTELALWKGVSQSLVGLAPRYQHTKAVIFT 437

QY 421 QNTGHWESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYEOTRGQSVSCLPL 480
DB 438 QVSRQWRKAKAVTGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEOTRGQSVSCLPL 497

QY 481 PRGQARWQCDVAVLYGQGPWGRFGAALTVLGVGVNGDKLTDVAIGAPGEEDNRGAVLYF 540
DB 498 PRGWR-RWCDVAVLYGQGPWGRFGAALTVLGVGVNGDKLTDVIGAPGEENRGAVLYF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQSHVLLRSQ 600
DB 557 HGVLGFSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTQDGLVDLAVGARGQVLLIATR 616
```

```
QY 601 PVLRLVKAIMFEFNPREVARNVFECNDQVWKGKAGEVVRVCLHVOKSTEDRLREGIOISVVT 660
DB 617 PVLWVGVSQFIPAEIPRSFAFECREQVWSEQTIVQSNICLYIDKRKNLLGSRDLQSSVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSTRRTQVTLGLTQTCTETLKLQLENCIEDPVSIVLRLNF 720
DB 677 LDIALDGRLSPRATQETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVTPITLRLNF 736

QY 721 SLVGTPLSAPGNLRPVLAEDAQRLFTALPFPEKNCNDNTICODDLSITTFMSLDCLVVG 780
DB 737 TLVGKPLLAFLPNLRPMLAADAQRYFTASLPEKNCADHICODNLGISFSPGLKSLVG 796

QY 781 GPREFNVTVVRNDGDSYRTQVTFEPPLDLSYRKYSTLONORSQSRMRLACSSASSTEV 840
DB 797 SNLELNAEVMVWMDGEDSYGTTITFHPAGLSYRYVAEGKQGGQLRLHLHITCDSAPVG-- 854

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
DB 855 SQGTWSTSCRINHILIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914

QY 901 QLELPVKYAVYVMTSHGVSTKYLNFAS-ENTSRVMQHOQYVSNLQORSIPISLVLVP 959
DB 915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORLDPVSINFWVP 974

QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIOCDI 1019
DB 975 VELNQEAVMMDVEVSHFPQNFSLRCSSEKIAPPASDFLAHQKPVLDSCSIAGCLFRCDV 1034

QY 1020 PFGIQIEFNATLGNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
DB 1035 PSRSVQELDFTLGNLSFGWVRQILQKKVSVSVASITFDTSVYSQLPQGEAFMRAQTT 1094

QY 1080 TKVEPFPVNPPLVLIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSE 1128
DB 1095 TVLEKYKVNPTPLIVGSSIGLLLLALITAVLYKGVFFKRYQKEMEE 1143

RESULT 6
ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=36111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sps and Sp3.";
RN J. Biol. Chem. 275:8959-8969 (2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=666289;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
RT beta 2-integrin alpha subunit.";
```



```
QY 481 PRGORARWOCDAVLYCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORVQWQCDVLYRGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGIGSPHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQGHVLLRSQ 600
Db 557 HGTSGSGIGSPHSORISAGSKLSPRLQYFGQALSGGQDLTMDGLVDLAVGARGQVLLRSL 616
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKQKGEAGEVRVCLHVQKSTRDRRLREQIQSVMVT 660
Db 617 PVLKVGVMARFSEVEKAVYRWEKPKSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVGLTGTCTETLKLQLPNCIEDPVPVSLVRLNLF 720
Db 675 FDLALDGRITSRATINENKPTLTKRKTGLGHCETLKLKLLPCDVEDVSPVLIHLNLF 734
QY 721 SLVGTPLSAGFNLRPVLAEDAQRLEFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 735 SLVREPIPSQNLRPVLAQSGDLFTASLPFEKNCQDGLCEGDLGVTLFSGLQTLTVG 794
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLQNRQSRQSWRLACSSASTEV 840
Db 795 SSLELVNIVTVWAGEDSYGVTVSLYYPAGLSHRVSGAQKPHQSAIRLACBTV-PTED 853
QY 841 SGALKSTCSINPIPEENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTKNTPE 900
Db 854 EG-LRSRCSVNHPIFHEGNGTFIVTFDVSYKATLGDRLMLKASSENKAKSSKATF 912
QY 901 QLELPVKYAVYVTVSHGVSTKYLNF-TASENTSRVMQHOYQVNSLQGRSLPISLVFLVP 959
Db 913 QLELPVKYAVYVTVSHGVSTKYLNF-TASENTSRVMQHOYQVNSLQGRSLPISLVFLVP 972
QY 960 VRLNQTVIMDRPQVTSSENLSSTCHKTERLPSSHDFLAELRKAPVWNCSTAVCORICQDI 1019
Db 973 VLLNGVAVVMDVMEAPSQSL--PCVSRKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1030
QY 1020 PFGIGIEEFENATLKNLSPDMYIKTSHNLLIYSTAEILLNDSVFTLLPGQAFVRSOTE 1079
Db 1031 PSFSVQEEFLTKLGNLSFQWVRETQKQVLVSVVAEITPDTSPQLPQEAQFMAQME 1090
QY 1080 TKVEPFEVFNPLFLVIGSSVGGLLLLALITAAALYKLGFFRQYKQKDMWSE 1128
Db 1091 MVLEEDVYNVNAIIMGSSVGALLLALITATLYKLGFFRKHYLEMLED 1139
RESULT 7
Q6KAS4
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:167-180(2004).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AK131133; BAD21383.1;
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR0002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3_
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR PRINTS; PR00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR SMART; SM00324; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS00234; VWA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDB4178 CRC64;

Query Match 56.7%; Score 3329.5; DB 2; Length 1188;
Best Local Similarity 56.7%; Pred. No. 3e-211;
Matches 649; Conservative 170; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
Db 39 FNLDSEKPTFHMDDGAEFGHSLVLYDSSWVVVGAPEIKATNQIGGLYKCGYHTGCEPI 98
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQCSTNTYVKGICFLFGSNLRQOPQK 120
Db 99 SLQVPPEAVNMSLGLSLAATNPWLLACGPTVHHTCRENIYLTGLCFLLSLSSFKQS-QN 157
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFISTVMEQLKSKTLFSLMQYSEEF 180
Db 158 FPTAQCECKPQDQIVFLIDGSGSISTDFEKLMDFKVAVMSQLQRPSTFSLMQFSDYF 217
QY 181 RHFTPEKPFQNNNPRSLKIPKLTQLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 218 RVHFTNNFISTSPSLDLSVRQLAGYTYTASAKHVITELEFTTGSGARQADATKLIVI 277
QY 241 TDGEKFGDPLGVEDVPIEADREGVIRYIVGVGDAPFSEKSRQELNTVASKPPRDHVFIQIN 300
Db 278 TDGRKQGDNLSDYSDVIPMAEASIIYYAIGVGKAFVNEHSKQBLKAIASMPSEHYVFSVE 337
QY 301 NEPAKLTIONLREKIFAETGTGSSSEFHEHMSQEGFSAATSTNGPLLSITVGSYDWAG 360
Db 338 NFDALDKIENQLKKEIFAETGTTPSSSTFEHMSQEGFSAVFTPGPVLGAYGVSFSWSG 397
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRRVQSLVGLAPRYQHTGLVAMFR 420
Db 398 GAFLYFSNMRPTFINNSQENEDMRDAYLGYSTALAPWKGVHSLJLGPAPRQHTGKVIFT 457
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGQGVSVCP 480
Db 458 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRGSTDLVLIGVPHYEHTRGQGVSVCPM 517
QY 481 PRGORARWOCDAVLYCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 518 P-GVGRRWCGTTLHGEQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVYIF 576
QY 541 HGTSGSGIGSPHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQGHVLLRSQ 600
Db 577 HGTSGSGIGSPHSORISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLAVGSKGVLLRTR 636
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKQKGEAGEVRVCLHVQKSTRDRRLREQIQSVMVT 660
Db 637 PILRVSPVTHFTPAETISRSVFECQVAPQETLSDATVCLIHESPKTQL--GDLRSTVT 694
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVGLTGTCTETLKLQLPNCIEDPVPVSLVRLNLF 720
Db 695 FDLALDGRITSRATINENKPTLTKRKTGLGHCETLKLKLLPCVAVESVTITLKLNF 754
QY 721 SLVGTPLSAGFNLRPVLAEDAQRLEFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 755 SLVGVPEISSLQNLQPLAVDDQTYFTASLPFEKNCADHICQDDLSVVFPGFPLKTLVVG 814
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLQ-----NQRSQSWR 829
Db 815 SDLELVNIVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLQ-----NQRSQSWR 874
```



```

QY 830 LACESASSTVSGALKSTSCSINHPIFPENSEVTENITFDVDSKASIGNKLLKANVTSE 889
Db 875 LMCD--STPDRSQGLWSTSCSRHVIFRGSGQMTFLVTFDVSRAELGDRLLLRARVGE 932
QY 890 NNMPRTNKTFFOLELPVKYAVYVMVTVSHGVSTKYNFTASE-NTSRVMOHQYQVSNLQOR 948
Db 933 NNVPGETKTTFFOLELPVKYAVYTMISSHDQFTKYNLFTSEKETSVEHRFQVNNLQOR 992
QY 949 SLFISLVFLVPLRLNQTIVWDRPQVTFSENLSTCHTKERLPHSDPLAELRKAPVNC 1008
Db 993 DVPVSINFVWPIELKGEAVW-TVMVSHPOPLTQCYRNLKPTQDFPLTHMQSPVLDCS 1051
QY 1009 IAVCORIQCDIPFFGQIEEFNATLKNLSPDWIKTSHNHLIVSTAELFNDVSFTLLP 1068
Db 1052 IADCLHLRCDIPSLGILDELFLFKGNLSFGWISQTLQKKVLLLSAEITFTNTSVYSQLP 1111
QY 1069 GQAFVRSQTEKTEKVEPFEVNPFLPIVSGSVGGLLLALITAAALYKLGFFKQYKDMWSE 1128
Db 1112 GQEAFLRAQTKTVLEMYKVHNPVPLIVSGSVGGLLLALITAILYKAGFFKQYKEMLEE 1171

RESULT 8
BAD21383
ID BAD21383 PRELIMINARY; PRT; 1188 AA.
AC BAD21383
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kitano R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131133; BAD21383.1;
FT NON_TER
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.7%; Score 3329.5; DB 2; Length 1188;
Best Local Similarity 56.9%; Pred. No. 3e-211;
Matches 649; Conservative 170; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQGLSYQCDYSTGSCPEI 60
Db 39 FNLDAEKPTFHMFGAEGFHSVLQYDSSWVVGAPKEIKATNQIGGLYKGYHTGNCEPI 98
QY 61 RLQVPEAVNMISGLSLAATTSPQLACGPTVHQCSENTYVYKGLCFPLGSLNRQOPQK 120
Db 99 SLQVPEAVNMISGLSLAATNPSWLLACGPTVHHTCRENIYLTGLCFLSSSFQSG-QN 157
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPFRRAKEFTISTVMEQKKSKTLFSLMQYSEEF 180
Db 158 FPTAQECQKQDQDIIVFLIDGSGSISTDPEKMLDFKAVMSQORPSTFSLMQSDYF 217
QY 181 RHFTTFEFONNPNRSLIKPITQLIGRTHATGIRKRVRELFINNGARKNAFKILILI 240
Db 218 RVHFTFNFTSTSPSLSDSVRLRGYTYTASAIKHVITELFTTQSGARQDQATKVLIV 277
QY 241 TQCEKFGDPLGYEDVLEPADREGVIRVILGVGDAFSEKSRQELNIVASKPPRDHVFQIN 300
Db 278 TDRKQGDNLSDYSVIPMAEASIIIRYAGVGKAFYNESKQELKAIASMPHSHEYFVSVE 337
QY 301 NFALKTIQNQLREKIFAIEGTQTGSSSSPEHMSQEGFSAATSNGLPLSTVGSDYDAG 360

```

RESULT 9

```

ID ITAX MOUSE
IT ITAX MOUSE STANDARD; PRT; 1169 AA.
AC Q9QX4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```

RA Huang X., Gorski K., Tong C., Rattis F.-M., Teeng S.-Y., Pardoll D.,
 RA Tsuchiya H.;
 RL "Isolation of genes selectively expressed by dendritic cells.";
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis (By similarity).
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF211864; AAF23492.1; --
 DR HSP; P20702; IN3Y.
 DR MGD; MGI:196509; Itgax.
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.
 DR InterPro; IPR00413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWFA; 1.
 DR PRINTS; PRO1185; INTEGRINA.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SMO0327; VWFA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 19 Potential.
 FT CHAIN 20 1169 Integrin alpha-X.
 FT DOMAIN 20 1116 Extracellular (Potential).
 FT TRANSMEM 1117 1137 Potential.
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 ? FG-GAP 2.
 FT DOMAIN 152 330 VWFA.
 FT REPEAT ? 402 FG-GAP 3.
 FT REPEAT 403 454 FG-GAP 4.
 FT REPEAT 456 518 FG-GAP 5.
 FT REPEAT 519 577 FG-GAP 6.
 FT REPEAT 582 634 FG-GAP 7.
 FT CA_BIND 467 475 Potential.
 FT CA_BIND 531 539 Potential.
 FT CA_BIND 594 602 Potential.
 FT SITE 1140 1144 GFFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 656 711 By similarity.
 FT DISULFID 770 776 By similarity.
 FT DISULFID 858 873 By similarity.
 FT DISULFID 1007 1031 By similarity.
 FT DISULFID 1036 1041 By similarity.
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 267 267 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 949 949 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1059 1059 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1084 1084 N-linked (GlcNAc...) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;
 Query Match 56.5%; Score 3320.5; DB 1; Length 1169;
 Best Local Similarity 56.8%; Pred. No. 1.1e-210;
 Matches 647; Conservative 172; Mismatches 302; Indels 19; Gaps 7;
 QY 1 FNLDTENAMTFQBNARGFQSGVVLQSGRVVVGAPQBIIVAAQORGSLSYQCDYSTGSGCEPI 60
 DB 20 FNLDAEKLTHFMDGAFFGHVLSVLYQDSSVVVVGAPKEIKATNQIGGLYKGYHTGNCCEPI 79
 QY 61 RLQVPVAVNMSLGLSLAATSPQLLIACGPTVHQTCTSENTYVKGCLFPLGSLNRQQPQX 120
 DB 80 SLQVPPAVANISLGLSLAATNPSWLLACGPTVHHTCRENIYLTGLCFLLSSFPKQSQ-QN 138
 QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFERRAKERISTVWEOLKSKTKLPSLMQYSEEF 180
 DB 139 FPTAQOQCPKQDDIVFLIDGSGSISSTDEKMLDFVKAVMSQLQRETSRPSLMQFSDYF 198
 QY 181 RIHFTFKFQNNPNRSLIKPIQTQLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
 DB 199 RVHFTFNNFISTSSPLSLGSLVRLGYTYTASAIKHVITELFTTQSGARQADATKVLIVI 258
 QY 241 TDGEKGPDLGYEDVPEADREGVIRKIVGVDAFRSEKSRQELNLTVASKPPDRHVFQIN 300
 DB 259 TDRGRKQGDNLSDSVIPMAEAAAIIRYAIQVGVKAFYNEHKSQELKALIASMPHSHEYVSFE 318
 QY 301 NFEALKTIONQLREKIFAIEGTOTGSSSSFEHMSQSGFAAITNSGPLLSTVGSYDWAG 360
 DB 319 NFDALKDIEQLKEKIFAIEGTETPSSSTFELEMSQSGFAVTPDGFVLGAVGSPFWSG 378
 QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAIAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
 DB 379 GAFLYPSNMRTPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTKGVVFT 438
 QY 421 QNTGMESNANVGTQIGAVFGASLCSVDVDSNGSTDLVLIGAPHYETQRTGGQSVCP 480
 DB 439 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDSGLDVLIGVPHYETHTRTGGQSVCPM 498
 QY 481 PRGORARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540
 DB 499 P-GVGSRWHCCTLHGEQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVIIF 557
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLRLRSQ 600
 DB 558 HGASRQDIAPSPQISASQIPSRIOYFGQSLGGQDLTRDGLVDLAVGSKRVLRLRTR 617
 QY 601 PVLRVKAIMEFNPREVARNVPECNDQVVKKEAGEVRLVHVQKSTDRDLREGIQSVWT 660
 DB 618 PILRVSPVTHFTPAEISRSVFECQEQVAPEQTLLSDATVCLHIESPKTLQ--GDLRSTVT 675
 QY 661 YDLALDSGRPHSRAVENETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVSPTVLRLNF 720
 DB 676 FDIALDHGRSLSTRAIFKETKTRALTRVKTGLGNKHCSVKLLLPACVEDSVPTILRLNF 735
 QY 721 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPFEKXNCNDNICODDLSTITFSFMSDLCLVVG 780
 DB 736 SLVGVPISSLQNLQPLAVDDQTYETASLPEKXNCADHICQDDLSSVVFPGFDLKLTVVG 795
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLSTRKVSTLQ-----NQBSQBSWR 829
 DB 796 SDLELNVDVTVSNDDGDSYGTITVTLFYFVGLSPRRVAEGQVFLRKKEQDQWQRGQSLH 855
 QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889
 DB 856 LMCD--STPDRSQGLWSTSCSRHVIPRGSGSQMTFLVTVSPKAEGLDRLLLRARVGE 913
 QY 890 NNMPRTNKTEFOLELPVKYAVMVVTVSHGVSTKYLNFASE--NTSRVMQHQYQVSNLQQR 948
 DB 914 NNVPGTPTKTFQLELPVKYAVYTMISSHDQFTKYLNFSTSEKETSVEVHRFQVNNLQQR 973
 QY 949 SLPLSLVLPVRLNQIVWDRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCS 1008
 DB 974 DVPVSINFVPIELKGEAVW-TVMVSHQPNPLTQCYRNRUKPTQFDLLTHMQKSFVLDCS 1032

```

QY 1009 IAVCORIQDIPFFGIOEBFNATKGNLSPDWIKTSHNHLIVSTAELFNDVSFTLLP 1068
Db 1033 IADCLHLRCIDPGLGIDELDFILKGNLSFGWISQTLQKVKLLSBAEITENTSVSYQLP 1092
QY 1069 GQAFVRSQTETKEVPEFNPPLPLIVSGSVGLLILALITAAALYKLGFFKROKXDMGSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVSGSVGLLILALITAAALYKAGFFKQYKEMLEE 1152

RESULT 10
ITAD RAT STANDARD; PRT; 1161 AA.
AC Q9QVE7:
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., Vandervieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC -!- VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWFA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.

```

```

FT REPEAT 88 ?
FT DOMAIN 152 334 VWFA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 456 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GFFKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.2%; Score 3244.5; DB 1; Length 1161;
Best Local Similarity 57.6%; Pred. No. 1.2e-205;
Matches 649; Conservative 163; Mismatches 302; Indels 13; Gaps 9;

QY 2 NLDNTAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYGTGSCPEIR 61
Db 21 NLDVEEPIVREDAAASFGTVVQFGSRLVGCAPLEAVAVNQTGRLYDCAPATMCOPIV 80
QY 62 LQVPVEAVNMSLGLSLAATTSPPQLLACGTPVHQTCSENYYVKGCLTFPGSMLRQOPKF 121
Db 81 LRSPLAEAVNMSLGLSLVATNNAQLLACGTPAQACVKMAYKGSCLLIGSSL-QFIQAV 139
QY 122 PEALRGCEQEDSDIAPLVDSGSIIPHPRRAKEFIISTVMEQKKSKTSLFSLMOYSEER 181
Db 140 PASMECEPRQEMDIAPLVDSGSIINORDFAQMKRVKALMGEPASTSTLFLSLMOYSNLK 199
QY 182 IHFTKEFQNNPNRSLIKPITQLLGRTHATGIRKVRVRELFNITNGARKNAFKILILIT 241
Db 200 THFTTEFKNILLDPQSLVDPIVQLQGLIYATGIRTWHEELFHSKNGSRKSAKILLVIT 259
QY 242 DGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNIVASKPPRDHVPQINN 301
Db 260 DGQKYRDPLEYSDDVIPAADKAGIIRYAIVGVGDAFQEPALKEINTIGSAPPQDHVPKVG 319
QY 302 PEALATIQNLREKIFALETGTGSSSFEHMSOGFSAAITNSGPIILSTVGSYDWAGG 361
Db 320 FAALRSIQRLQEKIFALETGTQSSRSFQHEMSQEGFSALTSDGPIGAVGSFWSGG 379
QY 362 VFLYTSKSKSTFINMTRVDSMDNDAYGYAAAILLRNRVQSILVGLAPRYQHIHLVAMERQ 421
Db 380 AFLYPNTPEFIFNNSENQVMDRSDYLGSTAVAFWKGVHSLILGAPRHQHTGKVIFITQ 439
QY 442 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYTRGQGVSVCLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSDDLVLIGAPHYYEYTRGQGVSVFPVP 499
QY 482 RCQARWQCDVAVLYCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYPH 541
Db 500 -GVRGRWQCEATLHCEQHPWGRFGVALTVLGDVNGDNLADVAIGAPGEESRGAIVYFH 558
QY 542 GTSGGIGISPHSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQP 601
Db 559 GASRLINPSPQRTVSGSLSLRLQYFGQSLGGQDLTQDGLVDLAVGAQGHVLLRSLP 618

```

QY	602	VLKVKALMEENPREVARNVFECDNDVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVTV	661
Db	619	LLKVELSIRFAPMEVAKAVYQCHERTPTVLEAGEATVCLVHKGSPDLL--GNVQGSVRY	676
QY	662	DLALDGRPHSRVFNFTKNSRQTOVLGLTQTCETLKLQPLNCIEDPSPVILRLNFS	721
Db	677	DLALDGRPLISRAIFDETCKNTLGRKTGLGDHCETVKLLPDCVEDAVSPIILRLNFS	736
QY	722	LVTGTPISAFGNLRPVLAEADQRLFTALPPEKNCNGNDNICODDLSTIFSPMSLDCLVVG	781
Db	737	LVRDSASP--RNLHPVLAVGSDHITASLPPEKNCQELLCEGLIGSFNFSGLQLVVG	795
QY	782	PREFNVTVTVRNDCEDSYRTQVTFPPDLDSYKRVSTLQORSQSWRLACESASSDEV	841
Db	796	SPELTVTVTVMNEGSDYGLVFKFYYPAGLSYRRTGTQ-QPHQYPLRLACEAPAAQED	854
QY	842	GALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEFQ	901
Db	855	--LRSSCSINHPIFREGAKTFTMTTFDVSYKAFGLGRLLLRKASSENKPDNTKTAQ	912
QY	902	LELPVKYAVVTVTSHGVSTKYNLFTASENTSR-VMOHQYQVSNLQORSPLISLVELVPV	960
Db	913	LELPVKYVTVTLRSQEDSTNHNFSSSHGRQEAARHYRVNLSPLKLAHVNFVWPV	972
QY	961	RLMOTVLDPRPQVTFSENLS--TCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDI	1019
Db	973	LINGVAWD---VTLSSPAQGVSCVQMKPPQDPDFLTQIRRSVLDCSIADCLHFRCDI	1029
QY	1020	PFGIQEEFNATLKNISFDWYIKTSHNHLIVSTABILFNDVFTLLPQGGAFVRSQTE	1079
Db	1030	PSLDIQDELFIILRNLNLSFGVWSQTLQEKVLLNSEAITFTDTSVYSLQPGAEFLRAQVE	1089
QY	1080	TKVEPEVFNPLPLIVGSSVGGILLIITALIYKLGPFKPKROYKDM 1126	
Db	1090	TILEYVYVYBPIFLVAGSSVGGILLIITWLYKLGFFKPKROYKEM 1136	
RESULT 11			
ID	ITAL_BOVIN	STANDARD;	PRT; 1165 AA.
AC	P61625;		
DT	05-JUL-2004 (Rel. 44, Created)		
DT	05-JUL-2004 (Rel. 44, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1		
DE	alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha		
DE	chain) (CD11a).		
GN	Names:ITGAL;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;		
RA	Fett T., Zechinon L., Baise E., Deamecht D.;		
RT	"The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,		
RT	characterisation and comparison with the human and murine		
RT	glycoproteins.";		
RL	Gene 325:97-101(2004).		
CC	-I- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,		
CC	ICAM3 and ICAM4. It is involved in a variety of immune phenomena		
CC	including leukocyte-endothelial cell interaction, cytotoxic T-cell		
CC	mediated killing, and antibody dependent killing by granulocytes		
CC	and monocytes (By similarity).		
CC	-I- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L		
CC	associates with beta-2 (By similarity).		
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).		
CC	-I- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins		
CC	with I-domains do not undergo protease cleavage.		
CC	-I- SIMILARITY: Belongs to the integrin alpha chain family.		
CC	-I- SIMILARITY: Contains 7 FG-GAP repeats.		
CC	-I- SIMILARITY: Contains 1 VWFA domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	EMBL; AY267467; AAP94035.1; -.		
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1.		
DR	PROSITE; PS00234; VWFA; 1.		
KW	Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;		
KW	Repeat; Signal; Transmembrane.		
FT	SIGNAL	23	Potential.
FT	CHAIN	24	Integrin alpha-L.
FT	DOMAIN	24	Extracellular (Potential).
FT	TRANSMEM	1085	Potential.
FT	DOMAIN	1106	Cytoplasmic (Potential).
FT	REPEAT	40	FG-GAP 1.
FT	REPEAT	89	FG-GAP 2.
FT	REPEAT	153	VWFA.
FT	REPEAT	347	FG-GAP 3.
FT	REPEAT	398	FG-GAP 4.
FT	REPEAT	454	FG-GAP 5.
FT	REPEAT	515	FG-GAP 6.
FT	REPEAT	575	FG-GAP 7.
FT	CA BIND	465	Potential.
FT	CA BIND	527	Potential.
FT	CA BIND	587	Potential.
FT	SITE	1111	GFPR motif.
FT	DISULFID	71	By similarity.
FT	DISULFID	108	By similarity.
FT	DISULFID	650	By similarity.
FT	DISULFID	768	By similarity.
FT	DISULFID	842	By similarity.
FT	DISULFID	994	By similarity.
FT	DISULFID	1017	By similarity.
FT	CARBOHYD	33	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	86	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	185	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	646	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	667	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	723	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	859	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	894	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	929	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1056	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1067	N-linked (GLCNAC. .) (Potential).
FT	SEQUENCE	1165	AA; 128725 MW; DAEB3A3F1E1463CB CRC64;
QY	1	FNLDTENAMTFQENARG--FGQSVVLQGSRRVVVGVAPQEIIVAAQNRSLYQCYSTGSC	58
Db	24	YNLDRVHVQNFSPFLAGRHFGRVQLQV-GNVVVVGVAPSE---GNSMNLVYQCFETGDC	79
QY	59	PIRLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVKGLCFPLGFSNLRQ	118
Db	80	PVTLSS--SNVTSKYLGWTLATDFTSDNLACDPLGSLRTCDQNIYLSGLCVLIHENLR	137
QY	119	OKPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTISTVMEQLKKSKTILSLMOYSE	178
Db	138	LQHPGQYQECIKGNVDLVFLFDGMSLQDQEFKIVDFMFKDKVMSKLSNSYQFAAVQ	197
QY	179	EERIHTFEKFNQNPNSRIKPIITOLGRTHTATGIRKVVRELFNITNGARKQAPKILI	238
Db	198	YRTEFTFLDIYRKOPDALLAGVKHRLNTNFGAINVYAKEVFRFDLGRPDATKVL	257

```
QY 239 LIIDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELINTVASKPRDRHVQ 298
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 258 IITDGEATD-----EHNIDAA--KDIIRYIIGIKNEFKTESQBALHQFASKPVEEFVKI 310
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 299 INNPEALKTTQONUREKIFAIEGTQTGSSSFHEMSQEGFSAAITSGNPLLTSTVGSYDW 358
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 311 LDTPEKLKDLFTLEQKIKYIVIEGTSKQDLTSFNMELSSSGISADLSSEGHVGVGAKDW 370
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 359 AGGVF-LYTSKEKSTFNMTVRSDMDNDAVLGYAAA-IILNRVQSLVIGAPYQHILGV 416
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 371 AGGLDLKADLKSTTFVGNELTVESRAGYLGYVTWLPGRGMSLLATGAPYQHVRGV 430
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQ 474
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 431 LLFQPKRGKGPWSQIQEIDGIGQSYFGELCGVDVDRDGETELLIAAPLYYGEQGG 490
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 475 VSVCLPRGQARWQDAVLIGEQQGWGRFGAALTVLGVNGVNGDKLTVAIGAPGEDNR 534
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 491 VFYI--QKIQLEFQWVSELQGETYPLGRFGAIAALTIDINGDELTDVAVGAPLEB--Q 545
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 535 GAVYLPHTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAGCHV 594
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 546 GAVYFNGQGG-GLSPSPSRIEQTQMFSGIOMFGRSIHGKDLGGDLADVAVGAGQV 604
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 595 LLRSQFVLRVKAIMFENPREARNFEPCNDQVVKGKEAG-EVRVCLHVKQ--STRDLR 651
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 605 IVLSPRPVDDIITSVSESPAEIPVHEVECSYSTSNQKKEGVNLTVCFOVKSLIST---F 660
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 652 EGOIQSVVTDLALDSGRPHRAVFNETKNSRTRQTVGLTQTCETFLKQLPNCIEDPV 711
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 661 QGHLVANLTYTLQDGHRTSRGLFPGGKHLGNTAVTPV-KSCFVFWFHFPCIQDLI 719
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 712 SPVLRINFLS---VGPFLS--AFGNLRPLVLAQAQRLFTALPPEKNCNDNICODDLS 766
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 720 SPINVSLSVLWEEGTDPDRALDRDIPILKPSHLETKEIPPEKNCNCKNCADLK 779
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 767 ITFSFMSLCLVVGPPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQONORSQ 826
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 780 LAFSDMRSKILRITPSASLSVRLTLNTAEDA-WVQVTLSPFQGLSFRKVEIL---KPHS 836
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 827 SWELACESASTVSGALKSTSCSINHPIIPENSEVTFTITFDVDSKASLGKLLKANY 886
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 837 HVPVGCLELPEEAHVHS-RALSCNVSPSPFGEDSMVDIQVMFNTLOGSGWDFIELQANV 895
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 887 TS-----ENNMPTNKTEFOLELPVKYAVVTVTSCHVSKYLNFTASENTRVMQH 938
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 896 SCNEDSSLLLEDNSATIS-----IPVMYPIINVLTQOENSTLYISFTPKSPIHHVKH 948
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 939 QYQV-----SNLQORSLPISLVFLVPLVRLNQTVI---WD---RPQVTFSS-ENLSSTCHKE 987
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 949 IVQVRIQPSNYDNP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESPSDEAE 1007
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 988 RUPSHSDFLAELKAPVNCSTIACVQRIQCDIPFGIQBEFNATLKNLSFDWYIKTSNN 1047
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1008 -----SCSFGT--EPRCPIDF--ROEILVQVNGMVELRGITIKAS-S 1043
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1048 HLLIVSTAEILNDSVFTLLPGGAFVRSTQTEKVEFEVNPENPLPLIVGSSVGLLALL 1107
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1044 MSLCSLSLAISFNSKXHLHYGRNASM-AQVVMKVDLVYEKEMLYLVLSGIGGLLFL 1102
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1108 ITAALYKLGFFKRYQKDM-----SRGPPGPAEQ 1137
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1103 IFITALYKVGFFKRLKKEKMEANVDASSETPGEDAGQPELEKE 1144
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 12

```
ID ITAL HUMAN STANDARD; PRT; 1170 AA.
AC P20701: 043746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
```

```
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=ITGAL; Synonyms=CD11A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.B., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036087; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain from
RT the CD11a/CD18 integrin.";
RL Structure 4:931-942(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weiss-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT CD11a I-domain.";
RL J. Mol. Biol. 292:1-9(1999).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Leukocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-CAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Y00796; CAA68747.1; --
 DR EMBL; AC002310; AAC31672.1; --
 DR PIR; S03308; S03308.
 DR PDB; 1CQP; X-ray; A/B=153-334.
 DR PDB; 1DGO; NMR; A=149-336.
 DR PDB; 1LFA; X-ray; A/B=--
 DR PDB; 1MJN; X-ray; A=153-331.
 DR PDB; 1MQ8; X-ray; B/D=155-331.
 DR PDB; 1MQ9; X-ray; A=152-330.
 DR PDB; 1MQA; X-ray; A=152-330.
 DR PDB; 1ZON; X-ray; @=150-336.
 DR PDB; 1ZOO; X-ray; A/B=150-336.
 DR PDB; 1ZOP; X-ray; A/B=150-336.
 DR Genew; HGNC:6148; ITGAL.
 DR MIM; 153370; --
 DR GO; GO:0008305; C:integrin complex; TAS.
 DR GO; GO:0006928; P:cell motility; TAS.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRIN.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA_1.
 DR PROSITE; PS00242; INTEGRIN ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 DR 3D-structure; Alternative splicing; Calcium; Cell adhesion;
 KW Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
 KW Receptor; Repeat; Signal; Transmembrane.
 FT SIGNAL 1 25
 FT CHAIN 26 1170 Integrin alpha-L.
 FT DOMAIN 26 1090 Extracellular (Potential).
 FT TRANSMEM 1091 1111 Potential.
 FT DOMAIN 1112 1170 Cytoplasmic (Potential).
 FT REPEAT 42 91 FG-GAP 1.
 FT REPEAT 92 149 FG-GAP 2.
 FT DOMAIN 156 327 VWFA.
 FT REPEAT 350 400 FG-GAP 3.
 FT REPEAT 401 455 FG-GAP 4.
 FT REPEAT 457 516 FG-GAP 5.
 FT REPEAT 518 575 FG-GAP 6.
 FT REPEAT 578 630 FG-GAP 7.
 FT CA_BIND 468 476 Potential.
 FT CA_BIND 530 538 Potential.
 FT CA_BIND 590 598 Potential.
 FT SITE 1115 1119 GFFXR motif.
 FT DISULFID 73 80 By similarity.
 FT DISULFID 111 129 By similarity.
 FT DISULFID 653 707 By similarity.
 FT DISULFID 771 777 By similarity.
 FT DISULFID 845 861 By similarity.
 FT DISULFID 998 1013 By similarity.
 FT DISULFID 1021 1052 By similarity.
 FT CARBOHYD 65 65 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 649 649 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 670 670 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 726 726 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 730 730 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 862 862 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 885 885 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 897 897 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1050 1060 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1071 1071 N-linked (GlcNAc...) (Potential).

Q -> OGVHGLVEMQTSKQILCRPAGDAEHTVGAQEGELPC
 PMGVSEAFRDNIRAGPCR (in isoform 2).
 /FTId=vsp_002738.
 R -> W (in Ref. 1 and 2).
 Y -> I (in Ref. 2).

FT	VARSPPLIC	954	954
FT	CONFLICT	214	214
FT	CONFLICT	660	660
FT	STRAND	155	162
FT	STRAND	164	164
FT	TURN	165	166
FT	HELIX	169	185
FT	TURN	186	198
FT	STRAND	191	198
FT	STRAND	202	206
FT	HELIX	208	214
FT	TURN	217	221
FT	TURN	222	223
FT	STRAND	229	229
FT	HELIX	233	243
FT	TURN	244	244
FT	HELIX	247	249
FT	TURN	250	250
FT	TURN	253	254
FT	STRAND	256	263
FT	TURN	274	276
FT	TURN	277	278
FT	STRAND	280	286
FT	HELIX	288	290
FT	HELIX	293	297
FT	TURN	298	299
FT	HELIX	300	302
FT	STRAND	307	310
FT	STRAND	311	314
FT	HELIX	317	319
FT	HELIX	323	330
FT	SEQUENCE	1170 AA;	128819 MW; 39A7AF92EF286FC0 CRC64;
FT	SEQUENCE	1170 AA;	128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.5%; Score 1557.5; DB 1; Length 1170;
 Best Local Similarity 34.7%; Pred. No. 7.1e-94;
 Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

Qy	1	FNLDTENAMTFQ--ENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLVQCDYSTGSC	58
Db	26	YNLDVRGARSFPSPRAGRHFGYRVLQV--GNGVIVGAPGE---GNSTGSLVQCDYSTGSC	81
Qy	59	PIRLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVVKGLCFLFGSNLR---	115
Db	82	PVTLR--GSNYTSKYLGMTLATDPTDGSILACDGLSRTCDQNTYLSGLCYLFRQNLQGM	140
Qy	116	-QQPKFPEALRGCPQSDSIAFLVDGSGSIIPHDFRRAKEFIISTVMEQLKKSKTLPSLM	174
Db	141	LQGRPGFOECIKG---NVDLVFLFDGSMISLQPDPEFQKILDFMKDVMKLSNTSYQFAAV	196
Qy	175	OYSEFRIHFTFEFQNNPNSRLIKPTQLLGHRTHTATGIRKVVYVRELFNITNGARKNAF	234
Db	197	QFSTSYKTEFDSDYVVRKOPDALLKHVGMLLLTNTFGAINVYVATEVFRRELGARPDAT	256
Qy	235	KILITIDGKFGPLGVEDVPEADREGVIRYVIGVGADAPRSEKSEKRELNTVASKPRD	294
Db	257	KVLIIITDGE--ATDSGNIDAADK-----IIRYIIGIKGHEQTKESQETLHKFKSPASE	309
Qy	295	HVFQINNFEALKTIONLREKIPIAETGTQTSSSPHEMSQEGFSAAITNSGILLSTVG	354
Db	310	FKVILDTFEKLKDLFTLEQKKIYVIEGTSKQDLTSFNMLSSSGISADLSRGHVAVGAVG	369
Qy	355	SYDNAGGVF-LYTSKEKSTFINMTFVSDMDNDAYLGYAAA--IIIRNRVQSVLVLGAPRYOH	412
Db	370	AKDWAGGFLDLKADLQDDTFIENEPLETVFVAGVILYTVTWLPSRQKTSLLASGAPRYOH	429
Qy	413	IGLVAMFR--QNTGMWESNANVKGTQICAYGASLCSVDVDSNGSTDLVLICAPHYYQQT	470
Db	430	MGRVLLFQEPQGGHWSQVQTIHGTQIGSYFGGELCGVDVDDQDGETELLIGCAPLFYGEQ	489
Qy	471	RGQVSVCPPLPRGORARWQCDAV--LYGECQCPWGRFGAALTGLVDVNGDKLTVAIGAP	528

```

Db 490 RGRVFYIY-----QRRQLGPEEVSELODQPGYPLGRGPEAITALTIDINGDLVDVAVGAP 544
QY 529 GEEDNRGAVLYEGTSGSISPSHSORLAGSKLSPRLQYEGQSLSGGQDLTMGLVDLTV 588
Db 545 LEE--QNAVIFNGRHG-GLSPQPSQRIGTQVLGSGIOWFGRSIHGVKDLGGGLADVAV 601
QY 589 GAQSHVLLRSQPYLRVKAIMEFNPREVARNVFECNDQVY-KGKEAGEVRVCLHVKQKSTR 647
Db 602 GAESQMVLSRPPVDMVMTLMSFPAEIPVHEVECSYSTSNKMGEGVNIICQI-KSLY 660
QY 648 DRLREGIQSVVYDIALDSGRHSRAVFNKTNKSTRQVGLGTQCTETLKLQLPNCI 707
Db 661 PQF-QGRLVANLTYTLQDGHRTRRGLFPGGRHELRLNAVT-TSMSCDFFSHPVGV 718
QY 708 EDVPSVILRLNLSL---VGTPLS--AFGN-----LRPVLAEDAQRLLFTALFPFEKNCGN 757
Db 719 QDLISPINVLSNLSLMEEGTGRDQRAQGDIPILRPSLSHSTWEI-----PFEXKCGE 773
QY 758 DNICODDLSTITFSFMSLDICIWGGPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVS 817
Db 774 DKCEANLRVSPARSRLRLTAFASLSVELSLNLEEDAYVQLDLHPPLGSLPRKVE 833
QY 818 TLQNRORSWRLACES--ASSTEVGALKSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCPEELPEESRLSRL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887
QY 876 LGNKLLKANVTSENN---MPTNKTETFOLELPVKYAVVMVTSHGVSSTKYLNTASEN 931
Db 888 WGSVSELHANVTNNEDSLLSENSATTI---IPILYPINILIQDQEDSTLYYSFYPKGP 944
QY 932 TSKVMQHOYQV---SNLGORSPL-LSLVFLVVPVRLNQTIVDWRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKHMQVRIQPSIHHDNIPTLEAVGVPPQPPSEGPITHQMSVQMEPPV--PCHYED 1002
QY 987 -BELPHSD--FLAELRKAPVNCSTAVCQRIQCDIPFGIQIEFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAEPCLPGALFRCPVW-----FROEILVQVIGTLELVGEIE 1044
QY 1044 TSHNHLIYSTABILENDSVFTLLPGQAFVRSTQETKYEFPVFPNPLPLIVGSSVGGLL 1103
Db 1045 AS-SMFSLCSSISISNSKHFHYGSNASL-AQVVMKVDVVYEQMQLYLYLSGGIGLL 1102
QY 1104 LLAALTAALYKLGFFRRQYKMDMSEG-GPPGAP 1136
Db 1103 LLLLIIFVLYKGVFFRNLKEMKAGVNGIP 1136

```

RESULT 13

```

Q6TYB8
ID Q6TYB8 PRELIMINARY; PRT; 1166 AA.
AC Q6TYB8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
SEQUENCE FROM N.A.
RA Dileepan T., Thumkikat P., Kannan M.S., Maheswaran S.K.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -! SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -! SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AY382558; AAQ90015.1; -.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.

```

```

DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1166 AA, 128723 MW, 78808AFBA896C9DF CRC64;

```

Query Match 26.5%; Score 1557; DB 2; Length 1166;

Best Local Similarity 34.6%; Pred. No. 7.6e-94;

Matches 409; Conservative 198; Mismatches 469; Indels 106; Gaps 33;

QY 1 FNLDTENAMTQENARG--PGQSVVQLQGSRVVVGAPQEIIVAANQKSLYQCDYSTGSC 58

Db 25 YNDLVHVHVNQFSPPLAGRHFGYRVLVQV-GNGVVVGAPSE---GNSMGNLYQCQFETGDCL 80

QY 59 PIRLVQVPEAVNMSLGLSLAATTPPQLLACGPTVHTCSENTYVYKGLCFGLFSGNLRQOP 118

Db 81 PVTLS--SNVTSKYLQMTLATDPTSDNLACDCLSRCTQDQNYLSGLCVLIHENLRGPV 138

QY 119 QKPEALRGCPQSDIAFLVDGSGSIIIPHDFFRAKEFISTVMEQLKKTKTLFSLMQYSE 178

Db 139 LQHPGYQECIKGNVDLVELFDGMSLSLQDDEFEKIVDFMKDVWVKLSNLSYQFAAQFST 198

QY 179 ERIHETEKEFQNNPNSRLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILI 238

Db 199 YRTEETFDIYIKQKDPDALLAGVHKMRLTNTFGAINVYAKEVFRPDILGARDPAKVI 258

QY 239 LITDGKFGDPLGYEDVIPEADREGVIRVYIVGVDAFRSEKRGQELNTVASKPRDHVFQ 298

Db 259 IITDGK--PPTNTTLMRPKTSS---RSLLGIGKNFKTESQEAHQFASKVEFEVKI 311

QY 299 INNFEALKTIQNLREKIPAIETGTGSSSPHEMSQBFSAITSNGLSTVGSYDW 358

Db 312 LDTFEKLKDLFTLQKKIYVIEGTSKQDLSFNMELSSSGISADLSEGHGHWGAVGAKDW 371

QY 359 AGSVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRVQSLVIGAPRYOHILV 416

Db 372 AGGLDLKADLSSTVGNELTVESRAGLYGVTVTRLPFSRGTMSSLATCAPKYQHVGRV 431

QY 417 AMFRQ--NTGMWSEANANVGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGG 474

Db 432 LLFQPKRGKQPSQIEIDIGIIGSYFGGELCGVDVDRDGETELLIAIAPLYGEGRGR 491

QY 475 VSVCLPRQARWQCDVLYGEGQGWQRFGAALTVLGDVNGDKLTDVAIGAPGEDNR 534

Db 492 VFIIY---QKIQLEFQWVSELOGETGYPLGRFGAAIAALTDINGDELTDVAVGAPLEE--Q 546

QY 535 GAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMGLVDLTVGAGHV 594

Db 547 GAVYIFNGQQG--GLSPRPSQRIEGTQMFSGIQWFGRSIHGVKDLGGDLADVAVGAGQV 605

QY 595 LLLRSQPVLRVKAIMEFNPREVARNVFECNDQVWVGKEAG-EVRVCLHVQK--STRDLR 651

Db 606 IVLSRPPVDIITSVSFPAEIPVHEVECSYSTSNKMGEGVNLTVCFQVKSLSIT----F 661

QY 652 EQGISVVTYDIALDSGRHSRAVFNKTNKSTRQVGLGTQCTETLKLQLPNCIEDPV 711

Db 662 QGHLVANLTYTLQDGHRTSRGLPFGKHKLIGNTAVTPV-KSCFVFWFHFPICIQDLI 720

QY 712 SPILVRLNLSL---VGTPLS--AFGNLRPVLAEDAQRLLFTALFPFEKNCGNIDICDDLS 766

Db 721 SPINVSLSYSLMEEEGTGRDPRALDRDIPILKPSHLETKEIPFEKNCGEDKNCBADLK 780

QY 767 ITFSFMSLDCLVVGQPREFNVTVRNDGEDSYRTQVTFPPPLDLSYRKVSFTLQNRQR 826

Db 781 LAFSDMRSKILLRTPSASISVRLTIRNTAEDAYVQVTLSPFQGLSFRKVEL---KPHS 837

QY 827 SWRLACESASSTEVGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANV 886

Db 838 HVPVGCCELPEEAVVHS-RALSCNVSSPIFGSDSMVDIQVMFNTLQKSGWGFIELQANV 896

QY 887 TS-----ENNMPRTNKTBEQLELPVKYAVVMVTSHGVSSTKYLNTFASENTSRVQ 938


```

Db      897 SCNNEDSLLEDNSATTS-----IPVMTPIINVLTQKQENSTLYISFTPKSPRIHVKH 949
QY      939 QYQV-----SNLQORSPLISLVFLVPLVRLNQTVI---WD---RPQVTFs-ENLSSTCHTKE 987
Db      950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPNLESDEAE 1008
QY      988 RLPSSHDFLAELKAPVWNCIAVCQRIQCDIPFGIOEFNATLKNLSFDWIKTSHN 1047
Db      1009 -----SCSFGT--EFRCPIDF---RQILVQVNGMWELRGTIKAS-S 1044
QY      1048 HLLIVSTAELFENDSVFTLLPGOGAFVRSOTETKVEPEFVNPPLPLIVGSSVGLLALL 1107
Db      1045 MLSCSLAISFNSSKHFHLHGRNASM-AQVWVKVDLYVEKEMLYLYVLSIGIGLLLLFL 1103
QY      1108 ITAALYKLGFFKQYKDDMM-----SEGGPFGBFPQ 1137
Db      1104 IFIYALYKVGFFKRNLEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 14
AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY382558; AAQ90015.1; -
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DFF CRC64;

```

```

Query Match 26.5%; Score 1557; DB 2; Length 1166;
Best Local Similarity 34.6%; Pred. No. 7.6e-94;
Matches 409; Conservative 198; Mismatches 459; Indels 106; Gaps 33;

QY      1 FNDTENAMTFQENARG--FGQSVVLQGGSRVVGAFQEIIVAAQNRSLYQCDYSTGSCSE 58
Db      25 YNLDRVHVQNFSPFLAGRHFGRYVLQV-GNGVVVVGAPSE---GNSMGNLYQCQPETGDL 80
QY      59 PIRLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVVKGL-CFLFGSNLRQOP 118
Db      81 PVLUS--SNYTSKYLGMVLTADPTSDNLLACDGLSKTCDQNIYLSGLCYLIHNLGRPV 138
QY      119 QKPEALRGCPQEDSIAFLVDGSGSIIPHDFRAKEBFISTVMEQLKKSKTLFLSMOYSE 178
Db      139 LQHPGQECIKGNVDLVFLFDGSMVSLQODEFEKIVDFMKDKMVKLSNYSYQFAVQFST 198
QY      179 EFRHFTFKPQNNPRLSIKPTQLLGRTHATGKRVKRVRELFNLTNGARKNAFKILI 238
Db      199 YFRTEFTFLDYIKQKDPDALLAGVKHMLLTNTFGAINYVAKEVFRPDLGARPDATKVI 258
QY      239 LITDEKFGDPLGVEDVPEADREGVIRYVGVDAPFRSEKSRQELNTVASKPPRDHVQ 298
Db      259 IITDCK---PPTNTILMRPTSS-----RSLGIGKNFKTKESQALHQPASKPVEEFVKI 311
QY      299 INNFEALKTIONLREKIFAIEGTQGTSSSFEHMQEFGSAATNSNPLLTSTGVSYDW 358
Db      312 LDTFEKLDLTELQKIVYIEGTSKQDLTSFNMLSSSGISADLSEGHGVGVGAVGAKDW 371
QY      359 AGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRRVQSLVILGAPRVQHIGLV 416
Db      372 AGGFLLDKADLKSTFVGNEPLTVESRAGYLGIVTVTRPLPSRGTMSLLATGAPKYQHVGRV 431

```

```

QY      417 AMFRQ--NTGMBSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQ 474
Db      432 LLFQQPKRGGPMQIQEIDGIGISYFGGELCGVDVDRDGETETELLIAAPLYGEGQGR 491
QY      475 VSYCPILPRGORARWQCDAVLYGEGQPGWGFGAALTIVLGVNGDKLTDVAIGAPBEDNR 534
Db      492 VFIV---QKIQLEBFQVMVSELOQETGYPLGRFGAIAALTDINGDELTDVAVGAPLEF-Q 546
QY      535 GAYLPHGTSGSIGISPHSQRISAGSKLSPLOYFGOSLGGQDLTMDGLVDLTVGAOGHV 594
Db      547 GAYIIFNGQGG-GLSRPFSQRIEGTFWFGIOWFGSRSHGVKDLGGGLADVAVGAGQV 605
QY      595 LLRSOPVLKRVKAIMFENPREVARNVFECDQNVVKGKEAG-EVRVCLHVQK--STRDLRL 651
Db      606 IIVLSRPVDIITVSFSPAEIFVHEVCSYSTSNQKKEGVNLTVCQVKSLLIST----F 661
QY      652 EGQIQSVVTVYDLALDGRPHSRVAFNETKNSTRQTVLGLTOTCETLKLQNLNCIEDPV 711
Db      662 QGHLVANLTYTLQDGHRTSRGLFPGKHKLGNTAVTPV-KSCFVFWFHFHFIQIDLI 720
QY      712 SPIVLRNFSL---VGTELS--AFGNLRPLVAEDAQRLEFALTALPFEKNCNDNICDDLS 766
Db      721 SPINVSLSYSLWEEGPRDRALDRDIPILPKSPHLETKYKIFFEKNCGEDKNCEADLK 780
QY      767 ITFSFMSLDCLVVGGPREFNVTIVRNDGEDSYRTQVTFEFLDLSYRKVSTLQNRQSR 826
Db      781 LAFSDMRSKTLRLTPSNASLSVRLTLRNTAEDAVYVQVTLSPQGLSPFRKVEIL--KPHS 837
QY      827 SWRLACESASTEVSGALKSTCSINHIPPENSEVTFNITFDVDSKASLGNKLLKXNV 886
Db      838 HVPVGCBELEPEEAVWHS-RALSCNVSSPIFGESMVDIQVMFNTLQKSGWGFELQANV 896
QY      887 TS-----ENNMPRTNKTEFQLELPVYAVVMVVTSHGVSTKYLNFTASENSTRVMOH 938
Db      897 SCNNEDSLLEDNSATTS-----IPVMTPIINVLTQKQENSTLYISFTPKSPRIHVKH 949
QY      939 QYQV-----SNLQORSPLISLVFLVPLVRLNQTVI---WD---RPQVTFs-ENLSSTCHTKE 987
Db      950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPNLESDEAE 1008
QY      988 RLPSSHDFLAELKAPVWNCIAVCQRIQCDIPFGIOEFNATLKNLSFDWIKTSHN 1047
Db      1009 -----SCSFGT--EFRCPIDF---RQILVQVNGMWELRGTIKAS-S 1044
QY      1048 HLLIVSTAELFENDSVFTLLPGOGAFVRSOTETKVEPEFVNPPLPLIVGSSVGLLALL 1107
Db      1045 MLSCSLAISFNSSKHFHLHGRNASM-AQVWVKVDLYVEKEMLYLYVLSIGIGLLLLFL 1103
QY      1108 ITAALYKLGFFKQYKDDMM-----SEGGPFGBFPQ 1137
Db      1104 IFIYALYKVGFFKRNLEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 15
ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1 alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha chain) (CD11a).
DE Name=Itgal; Synonyms=Lfa-1;
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RA "Cloning of the murine lymphocyte function-associated molecule-1

```



```

Db 716 IQDLISPINVSLNFSLEEGSTPRDQGRAMQPIRPSIHTV-TKEIPFEKNGCDKKCE 774
QY 763 DDLISITFSFMSLCLVVGGP-----REFNVTVTRNDGEDSYRTQVTFEPDLDSYRKV 816
Db 775 ANLTLSPPARS-----GPLRLMSSASLAVETWLSNGEDAYWVRLLDLPFRGLSPRKV 827
QY 817 STLQQRSSQRWRLACHSASTEVSGAL-KSTSCSINHPIPPENSEVTENITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKTKCNVSPPIFKAGQEVSLQVMFNTLLNS 882
QY 876 LGNKLILKANVTSEN-NMPTNKTEFOLELPVKYAVYVVTSHGVSTKYLNFTASENISR 934
Db 883 WEDEFVELNGTVHCENENSSLOEDNSAATHIPVLPVNLTKQENSTLYISFTPKGPKTQ 942
QY 935 VMQHQQVQVNLGQSRSLPISLVLVPRINQTVIWRDQO---VTFSENLS--TCHTK 986
Db 943 QVCHVYQV-----RIQPSAYDHNMT-LEALVGVRPHSEDLITYTWSVQTDPLVITCHSE 996
QY 987 E-RLPSHSDFLAELRKAPVYVNCSTAVCQRIQCDIPFFGQIEEFNATLKGNLSPDWYIKTS 1045
Db 997 DLKRPSE---AEQPCLPV-----QFRCPIVF---RWEILIQTVGTVELSKEIKAS 1042
QY 1046 HNHLIIVSTABILFNDSVFTLLPGQGAFVRSQETKVEPEVPNPPLIIVGSSVGGLLLL 1105
Db 1043 -STLSLCSLSVSNSSKHFLYGSKA-SEAQLVAVKVDLIHEKEMLHVYVLSGIGGLVLL 1100
QY 1106 ALITAAALYKLGFFKRYQYKDM-SEGPPGAP 1136
Db 1101 FLIFLALYKVGFFKRNLERMEADGGVNGSP 1132

```

Search completed: November 9, 2004, 12:17:38
 Job time : 141.75 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 ENLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAE PQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_23Sep04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	2	AAW65090 Human Bet
2	5862	99.8	1153	3	AAW65090 Human Bet
3	5862	99.8	1153	5	AAU80252 Human int
4	5862	99.8	1153	5	ABG61469 Human Bet
5	5862	99.8	1153	5	AAO14428 Integrin
6	5862	99.8	1153	7	ADD25615 Binding d
7	5852	99.6	1153	2	AAO4136 Alpha sub
8	5846.5	99.5	1152	8	ADM99589 Human int
9	5839.5	99.4	1152	8	ADP12435 Protein e
10	3483	59.3	1163	8	ADP44061 Human CDI
11	3469	59.0	1163	8	ADN02004 Human inf
12	3469	59.0	1163	8	ADQ17510 Human sof
13	3460	58.9	1163	2	AAO7120 p150.95 a
14	3446	58.6	1163	2	AAW65091 Human Bet
15	3446	58.6	1163	3	AAW65091 Human CDI
16	3446	58.6	1163	5	ABG61470 Human Bet
17	3444	58.6	1163	6	ABU07406 Protein d
18	3444	58.6	1163	7	ADG32005 Human hom
19	3411	58.0	1161	2	AAW78166 Human bet
20	3411	58.0	1161	2	AAW23049 Human bet
21	3411	58.0	1161	2	AAW57491 Human bet
22	3411	58.0	1161	2	AAW65089 Human Bet
23	3411	58.0	1161	2	AAW72825 Human alp
24	3411	58.0	1161	2	AAW73342 Human alp
25	3411	58.0	1161	3	AAW73359 Human alp

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; protein; 1153 AA.
XX AC AAW65090;

DT 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX rheumatoid arthritis.

OS Homo sapiens.

PN US5728533-A.

XX 17-MAR-1998.

PD 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

PR 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

PI Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or

XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is

XX used to describe a method for identifying compounds that modulate the

XX interaction of the beta-integrin alpha-d subunit with a binding partner

XX of alpha-d which involves contacting an alpha-d polypeptide with an alpha

XX -d binding partner, one of which is immobilised and the other of which is

XX labelled, in the presence of a test compound, and determining if the

XX compound affects binding between the alpha-d polypeptide and alpha-d

XX binding partner, where the alpha-d polypeptide is alpha-d or its fragment

XX comprising the cytoplasmic, transmembrane or extracellular domain of

XX alpha-d. Compounds that modulate alpha-d binding could be used to treat

ABG61468 Human Bet
AAW23064 Human Bet
AAW65106 Human Bet
AAW72837 Human alp
AAW73343 Human alp
AAW73376 Human alp
ABG61485 Human Bet
ABG61485 Human Bet
AAW78169 Rat alpha
AAW23062 Rat beta
AAW60004 Rat alpha
AAW72824 Rat alpha
AAW607374 Rat alpha
ABG61483 Rat Beta2
AAW65104 Rat beta-
AAW73345 Rat alpha
AAW23061 Mouse bet
AAW60003 Mouse bet
AAW65103 Mouse alp
AAW72836 Mouse alp
AAW73347 Mouse alp

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX
SQ Sequence 1153 AA;

Query Match 99.8%; Score 5962; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFQSVVQLQGSRRVVVGAPQEIIVAAANGSLYQCDYSTGSCPEPI 60
Db 17 FNLDTENAMTQENARGFQSVVQLQGSRRVVVGAPQEIIVAAANGSLYQCDYSTGSCPEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGCTVTHTQCSNTYVKGCLFGLFSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATSPPOLLAGCTVTHTQCSNTYVKGCLFGLFSNLRQOPQK 136
Qy 121 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTFLSLMOYSEEF 180
Db 137 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTFLSLMOYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELENTNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELENTNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSEQLNTVAKPPRHHVFOIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSEQLNTVAKPPRHHVFOIN 316
Qy 301 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVDSMDNAYLGYAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVDSMDNAYLGYAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQSVUCLP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQSVUCLP 496
Qy 481 PRGORARWQCDVLYGEOQWPGREGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGORARWQCDVLYGEOQWPGREGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660
Db 617 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGTOTCETLKLQLPNCIEDPVSPIVLRNF 736
Qy 721 SLVGTPLSAFNGNLRVLAEDAQRLLTALPFPEKNCNDNIQDDLSITFSFMSDCLVVG 780
Db 737 SLVGTPLSAFNGNLRVLAEDAQRLLTALPFPEKNCNDNIQDDLSITFSFMSDCLVVG 796
Qy 781 GPRFNVTVVRNDEGDSYRQVTFEFPDLSYKRVSTLQNRORSORSELACESASTEV 840
Db 797 GPRFNVTVVRNDEGDSYRQVTFEFPDLSYKRVSTLQNRORSORSELACESASTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENMPRINTKEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENMPRINTKEF 916
Qy 901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLFISLVFLVPV 960
Db 917 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLFISLVFLVPV 976

Qy 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDDIP 1020
Db 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDDIP 1036
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELFNDVSFTLLPGQCAFVRSQDET 1080
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELFNDVSFTLLPGQCAFVRSQDET 1096
Qy 1081 KVEFEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSEGGPPGAPQ 1137
Db 1097 KVEFEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSEGGPPGAPQ 1153

RESULT 2
AAB07360
ID AAB07360 standard; protein; 1153 AA.
XX
AC AAB07360;
XX
AC AAB07360;
DT 17-JAN-2001 (first entry)
XX
DE Human CD11b protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; aschma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11b.
XX
OS Homo sapiens.
XX
FN WO200029446-A1.
PD
XX 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
DR WPI; 2000-387751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
XX injury sites.
PS Example 5; Fig 1; 270pp; English.
XX

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11b. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAB07359). The present sequence has approximately 60% identity to the protein sequence of alpha.d. The Alpha.d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, aschma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease

SQ	Sequence 1153 AA;	
Query Match	99.8%; Score 5862; DB 3; Length 1153;	
Best Local Similarity	99.3%; Pred. No. 0;	
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 FNLDTENAMTPOENARFGQSVWOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60	
Db	17 FNLDTENAMTPOENARFGQSVWOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76	
Qy	61 RLQVPVAVNMSLGLSLAATTPOLLACGTPVHTQTCSENTYVKGCLFCLFGSNLRQOPQK 120	
Db	77 RLQVPVAVNMSLGLSLAATTPOLLACGTPVHTQTCSENTYVKGCLFCLFGSNLRQOPQK 136	
Qy	121 FPEALRCQPEDSDIAFLVDGSGIIIPHDPRRAKEFTSTWMEQLKSKTFLPSLMQYSEEP 180	
Db	137 FPEALRCQPEDSDIAFLVDGSGIIIPHDPRRAKEFTSTWMEQLKSKTFLPSLMQYSEEP 196	
Qy	181 RIHFTFEFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240	
Db	197 RIHFTFEFQNNPNRSLIKPIITOLLGRTHATGIRKVVRELFINITGARKNAFKILIVI 256	
Qy	241 TDGEKFGDPIGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTVASKPPRDHVFQIN 300	
Db	257 TDGEKFGDPIGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTIASKPPRDHVFQVN 316	
Qy	301 NFALKTIQNLREKIFAIEGTQTGSSSSSEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360	
Db	317 NFALKTIQNLREKIFAIEGTQTGSSSSSEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 376	
Qy	361 GVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAILLRNVQSLVLGAPRYQHIGLVAMFR 420	
Db	377 GVFLYTSKEKSTFINMTRVDSMDNDVILGYAAAILLRNVQSLVLGAPRYQHIGLVAMFR 436	
Qy	421 QNTGMWESNANVKGTOIGYFAGSLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVYCP 480	
Db	437 QNTGMWESNANVKGTOIGYFAGSLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVYCP 496	
Qy	481 PRGORARWQCDVILYGGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 540	
Db	497 PRGORARWQCDVILYGGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 556	
Qy	541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600	
Db	557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 616	
Qy	601 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRCVLHVQKSTRDLRREGIQSVVT 660	
Db	617 PVLVRKAIMEFNPREVARNVFCNDQVVKGEAGEVRCVLHVQKSTRDLRREGIQSVVT 676	
Qy	661 YDLALDSGRPHSRVAVNETKNSRTOVGLTQTCETLKLQLPNCIEDPVSIVLRINF 720	
Db	677 YDLALDSGRPHSRVAVNETKNSRTOVGLTQTCETLKLQLPNCIEDPVSIVLRINF 736	
Qy	721 SLVGTPLSAFGLNRPVLAEDAQRALTALFPFEKNCNDNTCQDLSITFFSMSLDCLVWG 780	
Db	737 SLVGTPLSAFGLNRPVLAEDAQRALTALFPFEKNCNDNTCQDLSITFFSMSLDCLVWG 796	
Qy	781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLRACASSTEV 840	
Db	797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSLRACASSTEV 856	
Qy	841 SGALKSTSCSINHPFIPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTFF 900	
Db	857 SGALKSTSCSINHPFIPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTFF 916	
Qy	901 QLELPVKYAYVMVVTSHGVSTKYLNFNTASENTSRVMOHYOVSNLGRSPISLIVLPV 960	
Db	917 QLELPVKYAYVMVVTSHGVSTKYLNFNTASENTSRVMOHYOVSNLGRSPISLIVLPV 976	
Qy	961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAFVNCISIAVCQRIQCDIP 1020	
Db	977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAFVNCISIAVCQRIQCDIP 1036	

Qy	1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQAFVRSOTET 1080	
Db	1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQAFVRSOTET 1096	
Qy	1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDDMMSEGGPPGAEPQ 1137	
Db	1097 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDDMMSEGGPPGAEPQ 1153	
RESULT 3		
AAU80252		
ID	AAU80252 standard; protein; 1153 AA.	
XX	AAU80252;	
AC	15-JUL-2002 (first entry)	
DT	Human integrin 1 alpha-M subunit protein.	
DE		
DE		
KW	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;	
KW	inflammatory disease; autoimmune disorder; Crohn's disease;	
KW	human immunodeficiency virus; HIV; myocardial infarction;	
KW	Sjorgen's syndrome; rheumatoid arthritis.	
XX		
OS	Homo sapiens.	
XX		
PH	Key Location/Qualifiers	
FT	Misc-difference 499..500	
FT	/note="Encoded by GGG CAG AGG"	
XX		
XX	WC200218583-A2.	
PD	07-MAR-2002.	
XX		
XX	31-AUG-2001; 2001WO-US027227.	
XX		
PR	01-SEP-2000; 2000US-0229700P.	
XX		
XX	(BLOO-) CENT BLOOD RES INC.	
PI	Springer TA, Shimoaka M, Lu C;	
XX		
DR	WPI; 2002-382964/41.	
DR	N-PSDB; ABK50046.	
XX		
PT	Modified integrin-I or integrin I-like domain polypeptide useful as an	
PT	immunogen to produce antibodies specific to polypeptide, comprises a	
PT	disulfide bond such that polypeptide is stabilized in a desired	
PT	conformation.	
XX		
PS	Disclosure; Page 109-112; 112pp; English.	
XX		
CC	This invention relates to a modified integrin-I or integrin I-like domain	
CC	polypeptide comprising at least one disulfide bond so that the domain is	
CC	stabilized in a desired conformation. The polypeptide of the invention	
CC	may have antiinflammatory or immunosuppressive activities. The	
CC	polypeptides of the invention have an open conformation and are useful	
CC	immunogens to produce antibodies that selectively bind to integrin I-	
CC	domain; and for identifying a modulator of integrin activity or of	
CC	interaction of an integrin and a cognate ligand. The polypeptide of the	
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for	
CC	treating or preventing an integrin mediated disorder which is an	
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the	
CC	binding of an integrin to a cognate ligand such as Crohn's disease,	
CC	nephritis; human immunodeficiency virus (HIV), myocardial infarction,	
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic	
CC	composition comprising the peptide of the invention is useful for	
CC	treating an integrin mediated disorder in a subject. The polypeptides	
CC	and/or active or antigenic fragments are useful as reagents for diagnosis	
CC	of integrin-mediated disorders. The present sequence represents the human	
CC	integrin-1 alpha-M protein subunit used to generate the mutant	
CC	polypeptides of the invention	

SQ	Sequence 1153 AA;	
Query Match	99.8%; Score 5862; DB 5; Length 1153;	
Best Local Similarity	99.3%; Pred. No. 0;	
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;	
QY	1 FNLDTENAMTQENARGFQSVOLQGSRRVVGAPQEIIVAAQNSGLYQCDYSTGSCPEI 60	
DB	17 FNLDTENAMTQENARGFQSVOLQGSRRVVGAPQEIIVAAQNSGLYQCDYSTGSCPEI 76	
QY	61 RLQPVAVNMSLGLSLAATTPPOLACGPTVHTQTSENTYVKGCLFLFGSNLRQPOK 120	
DB	77 RLQPVAVNMSLGLSLAATTPPOLACGPTVHTQTSENTYVKGCLFLFGSNLRQPOK 136	
QY	121 FPEALRGCPQDSDI AFLVDGSGSI PHDFRAKEFTSTWEOQKSKTLPFLMOYSEEF 180	
DB	137 FPEALRGCPQDSDI AFLVDGSGSI PHDFRMKEFTSTWEOQKSKTLPFLMOYSEEF 196	
QY	181 RIHFTFKFQNNPRSLIKPI TQLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240	
DB	197 RIHFTFKFQNNPRSLIKPI TQLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256	
QY	241 TDGEKFGDPLGYEVIPEADREGVIRVIVGVGDAPRSEKSRQELNIVASKPPRDHVFQIN 300	
DB	257 TDGEKFGDPLGYEVIPEADREGVIRVIVGVGDAPRSEKSRQELNIVASKPPRDHVFQVN 316	
QY	301 NFEALKTIQNLREKI FAIEGTQTGSSSPHEMSQSGFAAITSNGPILLSVGSYDWAG 360	
DB	317 NFEALKTIQNLREKI FAIEGTQTGSSSPHEMSQSGFAAITSNGPILLSVGSYDWAG 376	
QY	361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420	
DB	377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436	
QY	421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVIGAPHYEYQTRGGQSVCP 480	
DB	437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVIGAPHYEYQTRGGQSVCP 496	
QY	481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 540	
DB	497 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNREGAVYLF 556	
QY	541 HGTSGSGISPSHSGRIAGSKLSPLQYFGOSLGGQDLTMDGLVLTGVAQGHVLLRSQ 600	
DB	557 HGTSGSGISPSHSGRIAGSKLSPLQYFGOSLGGQDLTMDGLVLTGVAQGHVLLRSQ 616	
QY	601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVVRVCHVQKSTRDRUREGQIQSVVT 660	
DB	617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVVRVCHVQKSTRDRUREGQIQSVVT 676	
QY	661 YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPTVLRLNF 720	
DB	677 YDLALDSGRPHSAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPTVLRLNF 736	
QY	721 SLVGTPLSAFGLNLRPLVAEDAQLFTALFPFEKNCNDNICQDDLSITFFSMLDCLVVG 780	
DB	737 SLVGTPLSAFGLNLRPLVAEDAQLFTALFPFEKNCNDNICQDDLSITFFSMLDCLVVG 796	
QY	781 GPREFNVTVVRNDGEDSYRTQTFPPPLDSLVRKYSTLQNSQRSWRILACSASSTEV 840	
DB	797 GPREFNVTVVRNDGEDSYRTQTFPPPLDSLVRKYSTLQNSQRSWRILACSASSTEV 856	
QY	841 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKXANTYSENMPRTNKTFF 900	
DB	857 SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGKLLKXANTYSENMPRTNKTFF 916	
QY	901 QLELPVKYAYVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960	
DB	917 QLELPVKYAYVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 976	
QY	961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVQRIQCDIP 1020	
DB	977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVQRIQCDIP 1036	
QY	1021 PFGIOEEFNATLKNLSFDWYIKTSHNHLILIVSTAELFNDVSFTLLPQGAFVRSQET 1080	
DB	1037 PFGIOEEFNATLKNLSFDWYIKTSHNHLILIVSTAELFNDVSFTLLPQGAFVRSQET 1096	
QY	1081 KVBPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPBPGEPO 1137	
DB	1097 KVBPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPBPGEPO 1153	
RESULT 5		
AAO14428		
ID	AAO14428 standard; protein; 1153 AA.	
AC	AAO14428;	
XX		
DT	03-MAY-2002 (first entry)	
XX		
DE	Integrin Mac-1 alpha subunit.	
XX		
KW	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
KW	open conformation; integrin related inflammatory disorder;	
KW	integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW	reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW	viral infection; cancer; gene therapy; vaccine;	
KW	bioactive agent screening.	
XX		
OS	Unidentified.	
XX		
PN	W0200204521-A2.	
XX		
PD	17-JAN-2002.	
XX		
PF	09-JUL-2001; 2001WO-US021805.	
XX		
PR	07-JUL-2000; 2000US-0216600P.	
XX		
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.	
PA	(BLOO-) CENT BLOOD RES.	
XX		
PI	Springer T;	
XX		
DR	WPI; 2002-148167/19.	
XX		
PT	New integrin I domain protein having alteration in at least 2	
PT	noncontiguous regions and exits in an open conformation, useful for	
PT	treating, preventing or suppressing inflammatory or immunological	
PT	disorders.	
XX		
PS	Example 1; Fig 1F; 90pp; English.	
XX		
CC	The invention comprises structurally biased variant integrin inserted (1)	
CC	domain proteins, wherein the alterations to the protein occur in at least	
CC	two noncontiguous regions. Specifically the variant integrin I domain	
CC	proteins are structurally biased to exist in the open conformation,	
CC	thereby altering the binding ability of the protein. The invention also	
CC	comprises nucleic acids encoding the variant integrin I domain proteins.	
CC	The integrin I domain proteins and nucleic acids are useful for treating,	
CC	preventing or suppressing integrin related inflammatory and immunological	
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain	
CC	proteins and nucleic acids can also be used for treating: ischaemia/	
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral	
CC	infection; and cancer. The variant integrin I domain nucleic acids and	
CC	proteins may be used in gene therapy, as vaccines and to screen for	
CC	bioactive agents. The present amino acid sequence represents the Mac-1	
CC	alpha subunit of integrin	
XX		
SQ	Sequence 1153 AA;	
Query Match	99.8%; Score 5862; DB 5; Length 1153;	
Best Local Similarity	99.3%; Pred. No. 0;	
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;	

QY 1 FNLDTENAMTFOENARGFGQSQVWOLQSGRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCBPI 60
Db |||||
QY 17 FNLDTENAMTFOENARGFGQSQVWOLQSGRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCBPI 76
Db |||||
QY 61 RLQVPVAVNMSLGLSLAAATSPQLLACGPTVHQTCSNTYVVKGLCFLFGSLNRQQPQK 120
Db |||||
QY 77 RLQVPVAVNMSLGLSLAAATSPQLLACGPTVHQTCSNTYVVKGLCFLFGSLNRQQPQK 136
Db |||||
QY 121 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQLKSKTSLFSLMQYSEBF 180
Db |||||
QY 137 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQLKSKTSLFSLMQYSEBF 196
Db |||||
QY 181 RIHTEKEFQNNPNSRLIKIPITOLGRTHATGIRKVVRELENTNGARKNAFKILILI 240
Db |||||
QY 197 RIHTEKEFQNNPNSRLIKIPITOLGRTHATGIRKVVRELENTNGARKNAFKILIVI 256
Db |||||
QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVFQIN 300
Db |||||
QY 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVFQIN 316
Db |||||
QY 301 NFEALKITQNLREKIFAIETGOTGSSSFHEMSQEGFSAATNSNGLPSTVGSYDWAG 360
Db |||||
QY 317 NFEALKITQNLREKIFAIETGOTGSSSFHEMSQEGFSAATNSNGLPSTVGSYDWAG 376
Db |||||
QY 361 GVELYTSKEKSTFINMTRVDSMDNDAYLGXAAIILNRNVSOLVGLGAPRYOHIGLVAMFR 420
Db |||||
QY 377 GVELYTSKEKSTFINMTRVDSMDNDAYLGXAAIILNRNVSOLVGLGAPRYOHIGLVAMFR 436
Db |||||
QY 421 QNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEQTGRGGVSVCP 480
Db |||||
QY 437 QNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEQTGRGGVSVCP 496
Db |||||
QY 481 PRGQARWQCDVAVLGEQGPWGRGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db |||||
QY 497 PRGQARWQCDVAVLGEQGPWGRGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556
Db |||||
QY 541 HGTSGSGISPHSQRISAGSLPRLOYFGQSLGGQDLTMGLVLTGVAQHVLILRSQ 600
Db |||||
QY 557 HGTSGSGISPHSQRISAGSLPRLOYFGQSLGGQDLTMGLVLTGVAQHVLILRSQ 616
Db |||||
QY 601 PVLRYKAIEMFNPREVARNFECDNVQVKGKEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db |||||
QY 617 PVLRYKAIEMFNPREVARNFECDNVQVKGKEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
Db |||||
QY 661 YDLALDSRPHSRAVFNETHKSTRTOVLGHTQTCETILKLOPNCIEDPVSPIVRLNF 720
Db |||||
QY 677 YDLALDSRPHSRAVFNETHKSTRTOVLGHTQTCETILKLOPNCIEDPVSPIVRLNF 736
Db |||||
QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALFPFKNKNGNDNICODDLSITFSFMSLDCLVVG 780
Db |||||
QY 737 SLVGTPLSAFNLRPVLAEDAQLFTALFPFKNKNGNDNICODDLSITFSFMSLDCLVVG 796
Db |||||
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLRYKRVSTLQNRQSRWPLACESASSTEV 840
Db |||||
QY 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLRYKRVSTLQNRQSRWPLACESASSTEV 856
Db |||||
QY 841 SGALKSTSCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSENMPNKNKEF 900
Db |||||
QY 857 SGALKSTSCSINHPIFPENSEVTFNITPDVDSKASLGNKLLKANVTSENMPNKNKEF 916
Db |||||
QY 901 QLELPVKVAVNMVTVSHGVSTKYLNFASNTSRVNHQYQVSNLQSRSLPISLFLVFPV 960
Db |||||
QY 917 QLELPVKVAVNMVTVSHGVSTKYLNFASNTSRVNHQYQVSNLQSRSLPISLFLVFPV 976
Db |||||
QY 961 RLNQTVIWRDPQVTFSENISSCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1020
Db |||||
QY 977 RLNQTVIWRDPQVTFSENISSCHTKERLPSHSDFLAELRKAPVNCSTAVCQRTQCDIP 1036
Db |||||
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAILFNDVSVFTLLPQCGAFVRSQTF 1080
Db |||||
QY 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAILFNDVSVFTLLPQCGAFVRSQTF 1096
Db |||||
QY 1081 KVEPPEVNPPLIVGSSVGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1137

Db 1097 KVEPPEVNPPLIVGSSVGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1153
|||
RESUT 6
ADD25615
ID ADD25615 standard; protein; 1153 AA.
XX ADD25615;
XX 15-JAN-2004 (first entry)
XX Binding domain-immunoglobulin fusion protein-associated protein #85.
XX Binding domain; immunoglobulin; fusion protein; cytostatic;
XX antiarthritic; immunosuppressive; antidiabetic; antithyroid;
XX neuroprotective; hinge region; immunoglobulin heavy chain;
XX CH2 constant region; CH3 constant region; IgG1;
XX antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;
XX malignant condition; B-cell disorder; melanoma; sarcoma;
XX rheumatoid arthritis; myasthenia gravis; Grave's disease;
XX type I diabetes mellitus; multiple sclerosis; autoimmune disease.
XX Unidentified.
OS US2003118592-A1.
XX 26-JUN-2003.
PD 25-JUL-2002; 2002US-00207655.
XX 17-JAN-2001; 2001US-0367358P.
PR 17-JAN-2002; 2002US-00053530.
PR 03-JUN-2002; 2002US-0385691P.
XX (GENE-) GENE-CRAFT INC.
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;
PI WPI; 2003-803137/75.
DR New binding domain-immunoglobulin fusion protein, useful for treating a
subject having or suspected of having a malignant condition or a B-cell
disorder, e.g. melanoma, Grave's disease or autoimmune disease.
XX Disclosure; SEQ ID NO 176; 157pp; English.
CC The invention relates to a binding domain-immunoglobulin fusion protein
comprising a binding domain polypeptide that is fused to an
immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain
CH2 constant region polypeptide that is fused to the hinge region
polypeptide, and an immunoglobulin heavy chain CH3 constant region
polypeptide that is fused to the CH2 constant region polypeptide. The
hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin
hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge
region polypeptide, derived from (a) having 3 or more cysteine residues;
where the mutated human IgG1 immunoglobulin hinge region polypeptide
contains 2 cysteine residues, where the first cysteine is not mutated; a
mutated human IgG1 immunoglobulin hinge region polypeptide, derived from
(a) having 3 or more cysteine residues, where the mutated human IgG1
immunoglobulin hinge region polypeptide contains no more than one
cysteine residue; and a mutated human IgG1 immunoglobulin hinge region
polypeptide, derived from (a) having 3 or more cysteine residues; where
the mutated human IgG1 immunoglobulin hinge region polypeptide contains
no cysteine residues. The binding domain-immunoglobulin fusion protein is
capable of at least one immunological activity comprising antibody
dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The
binding domain polypeptide is capable of specifically binding to an
antigen. Also included are an isolated polynucleotide encoding the
binding domain-immunoglobulin fusion protein, a recombinant expression
construct comprising the polynucleotide (operably linked to a promoter),
a host cell transformed or transfected with a recombinant expression
construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin
 CC fusion protein or polynucleotide and a carrier, and treating a subject
 CC having or suspected of having a malignant condition or a B-cell disorder.
 CC The binding domain-immunoglobulin fusion protein is useful for treating a
 CC subject having or suspected of having a malignant condition or a B-cell
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple
 CC sclerosis or autoimmune disease. The present sequence is a binding domain
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The
 CC sequence data for this patent formed part of the printed specification
 CC and is also available in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not
 CC identified the sequences in the printed specification by their SEQ ID
 CC number therefore none of the sequences can be explicitly identified.

XX
 SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 7; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVQLQGSRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 60
 DB 17 FNLDTENAMTFQENARGFGQSVQLQGSRVVVGAPQBIIVAAHQSGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOPK 120
 DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOPK 136
 QY 121 FPEALRGCPQEDSDIAFLPDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQXSEEF 180
 DB 137 FPEALRGCPQEDSDIAFLPDGSGSIIPHDFRRAKEFISTVMEQLKKSKTLFSLMQXSEEF 196
 QY 181 RIHFTKEFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
 DB 197 RIHFTKEFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
 QY 241 TDEKEFGDPLGYEDVPEADREGVIRYVIGVGAFAFRSEKSRQELNTVASKPPRDHVFQIN 300
 DB 257 TDEKEFGDPLGYEDVPEADREGVIRYVIGVGAFAFRSEKSRQELNTVASKPPRDHVFQVN 316
 QY 301 NFPEALTIQNLREKIFAIBGTQTGSSSSPEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
 DB 317 NFPEALTIQNLREKIFAIBGTQTGSSSSPEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
 QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGVAIAIILRNVRQSLVIGAPRYQHTGLVAMFR 420
 DB 377 GVFLYTSKEKSTFINTRVDSMDNDAYLGVAIAIILRNVRQSLVIGAPRYQHTGLVAMFR 436
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVEQTRGGQVSCPL 480
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYVEQTRGGQVSCPL 496
 QY 481 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
 DB 497 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
 QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQVFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQVFGQSLSGQDPLTMDGLVDLTVGAQGHVLLRSQ 616
 QY 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLVHVKSTRDRREGQIOSVVT 660
 DB 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLVHVKSTRDRREGQIOSVVT 676
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQPNCIEDPVSPIVLRINF 720
 DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQPNCIEDPVSPIVLRINF 736
 QY 721 SLVGTPLSARCNRPVLAEDAQLFTALFPFENKCGNDNICQDDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSARCNRPVLAEDAQLFTALFPFENKCGNDNICQDDLSITFSFMSLDCLVVG 796

QY 781 GPREENVTVRNDGEDSVRTQVTFEPFLDLSYRKUSTLONORSQSRWLACESASSTEV 840
 DB 797 GPREENVTVRNDGEDSVRTQVTFEPFLDLSYRKUSTLONORSQSRWLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFOVDSKASLGKLLKANVTSENMPRTNKTEF 900
 DB 857 SGALKSTSCSINHPIFPENSEVTFNITFOVDSKASLGKLLKANVTSENMPRTNKTEF 916
 QY 901 QLELPVKIAYVMVTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSILPISLVFLVPV 960
 DB 917 QLELPVKIAYVMVTSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSILPISLVFLVPV 976
 QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
 DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036
 QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTE 1080
 DB 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTE 1096
 QY 1081 KVEPFVEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
 DB 1097 KVEPFVEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 7
 AAR04136
 ID AAR04136 standard; protein; 1153 AA.
 XX AAR04136;
 AC AC
 XX 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 07-SEP-1990 (first entry)
 XX
 DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
 XX
 KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
 KW non-specific defence system; integrin gene superfamily.
 XX
 OS Synthetic.
 PH Key
 FT Region
 FT 1..116
 FT /label= signal_peptide
 FT Modified-site 86..88
 FT /label= putative N-glycosylation site
 FT Modified-site 240..242
 FT /label= putative N-glycosylation site
 FT Modified-site 391..393
 FT /label= putative N-glycosylation site
 FT Modified-site 469..471
 FT /label= putative N-glycosylation site
 FT Modified-site 693..695
 FT /label= putative N-glycosylation site
 FT Modified-site 697..699
 FT /label= putative N-glycosylation site
 FT Modified-site 735..737
 FT /label= putative N-glycosylation site
 FT Modified-site 802..804
 FT /label= putative N-glycosylation site
 FT Modified-site 881..883
 FT /label= putative N-glycosylation site
 FT Modified-site 901..903
 FT /label= putative N-glycosylation site
 FT Modified-site 912..914
 FT /label= putative N-glycosylation site
 FT Modified-site 941..943
 FT /label= putative N-glycosylation site
 FT Modified-site 947..949
 FT /label= putative N-glycosylation site
 FT Modified-site 979..981
 FT /label= putative N-glycosylation site

PT Modified-site 994..996
/label= putative N-glycosylation site
FT Modified-site 1022..1024
/label= putative N-glycosylation site
FT Modified-site 1045..1047
/label= putative N-glycosylation site
FT Modified-site 1051..1053
/label= putative N-glycosylation site
FT Modified-site 1076..1078
/label= putative N-glycosylation site
FT Region 1106..1134
/label= putative_transmembrane_region
XX
EP364690-A.
XX
XX 25-APR-1990.
XX
XX 17-AUG-1989; 89EP-00115159.
XX
XX 23-AUG-1988; 88US-00235353.
PR 09-MAR-1989; 89US-00321239.
XX
XX (DAND) DANA FARBER CANCER INST INC.
PA
XX
XX Springer TA, Corbi A;
XX WPI; 1990-125938/17.
DR N-PSDB; AAQ04043.
XX
XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
PT inflammation and viral infections, and in diagnosis.
PT
PT
PS Disclosure; Page ?; 3pp; English.
XX
XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.
CC recognition of and migration to sites of inflammation. It also attaches
CC to cellular substrates as part of this function making it useful in
CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
CC -MAR-2003 to correct PA field.)
CC
CC Revised record issued on 09-SEP-2004 : Correction to feature table key
CC and pages
XX
XX Sequence 1153 AA;
SQ
Query Match 99.6%; Score 5852; DB 2; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
QY 1 FNLDTENAMTQENARGFQSVVLOGSRVVVGAPOEIVAAQNRGSLVQCDYSTGSCBPI 60
DB 17 FNLDTENAMTQENARGFQSVVLOGSRVVVGAPOEIVAAQNRGSLVQCDYSTGSCBPI 76
QY 61 RLOVPEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFGLFGLNLRQQPQK 120
DB 77 RLOVPEAVNMSGLSLAATSPOLLACGPTVHTCSENTYVKGCLFGLFGLNLRQQPQK 136
QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFFRAKEFIITVMEMQLKSKTFLSLMOYSBEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFFRMKEFVSTVMEMQLKSKTFLSLMOYSBEF 196
QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRHTATGIRKVVRELFINITNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNSRLIKPTQLLGRHTATGIRKVVRELFINITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTIASKPPRDHFQIN 316
QY 301 NFEALKTIONQREKIFALEGHTQGTSSSFEHMSQEGHSAITNGPILLSVGYSDWAG 360
DB 317 NFEALKTIONQREKIFALEGHTQGTSSSFEHMSQEGHSAITNGPILLSVGYSDWAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLVGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLVGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQVSVCP 496
QY 481 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVYNGDKLTDVAIGARGEEDNAGVYLF 540
DB 497 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVYNGDKLTDVAIGARGEEDNAGVYLF 556
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAOCHVLLLRSQ 600
DB 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAOCHVLLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPEKKNCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPEKKNCGNDNICQDDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHIFPENSEVTNITFDVDSKASLGNKLLIKANVTSENWMPRTNKTEF 900
DB 857 SGALKSTSCSINHIFPENSEVTNITFDVDSKASLGNKLLIKANVTSENWMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTRVMQHOYQVSNLQORSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTRVMQHOYQVSNLQORSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPAPVNVNCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKPAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVSQDET 1080
DB 1037 PFGIQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVSQDET 1096
QY 1081 KVEPFEVNPPLTVGVSSVGGLLLLALITALYKLGFFPKROYKDMWSEGGPPGAEPPQ 1137
DB 1097 KVEPFEVNPPLTVGVSSVGGLLLLALITALYKLGFFPKROYKDMWSEGGPPGAEPPQ 1153
RESULT 8
ADM99589
ID ADM99589 standard; protein; 1152 AA.
XX
AC ADM99589;
XX
DT 17-JUN-2004 (first entry)
XX
DE Human integrin alphaM subunit precursor protein.
XX
XX integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;
KW osteopathic; cytostatic; immunosuppressive; antiinflammatory;
KW neuroprotective; antistickling; immunotherapy; inflammatory;
KW autoimmune disorder; thrombosis; cancer; osteoporosis;
KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
KW alphasM.
XX
OS Homo sapiens.
XX

FH Key Location/Qualifiers
 FT Misc-difference 965 /note= "Encoded by CCC"
 XX W02004007530-A2.
 XX 22-JAN-2004.
 XX 17-JUL-2003; 2003WO-US022301.
 XX 17-JUL-2002; 2002US-0396783P.
 PR 17-JUL-2002; 2002US-0396790P.
 PR 11-SEP-2002; 2002US-0410135P.
 XX
 XX (BLOO-) CENT BLOOD RES INC.
 XX
 XX Springer TA, Takagi J;
 XX WPI; 2004-122877/12.
 DR N-PSDB; ADM99588.
 XX
 XX Novel modified integrin protein having extracellular domains of integrin
 PT alpha and beta subunits or integrin alphaII and beta3 subunit, useful for
 PT treating integrin mediated disorders.
 XX
 XX Disclosure; SEQ ID NO 4; 232pp; English.
 XX
 XX The invention relates to a novel isolated or recombinant modified
 CC integrin protein having extracellular domains of integrin alpha and beta
 CC subunits where one of the subunits has one or more mutations, an altered
 CC surface feature or an amino acid substitution or internal deletion,
 CC extracellular domains of the integrin beta subunit that comprise a
 CC mutation that alters a non-cysteine residue to cysteine or extracellular
 CC domains of integrin alpha and beta subunits. The polypeptide of the
 CC invention demonstrates antiproliferative, thrombolytic, anticoagulant,
 CC osteoprotective, cytoskeletal, immunosuppressive, antiinflammatory,
 CC neuroprotective and antisticking activities and may be useful for
 CC immunotherapy in order to prevent or treat an integrin-mediated disorder
 CC such as an inflammatory disorder, an autoimmune disorder, thrombosis,
 CC cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple
 CC sclerosis. The current sequence is that of the human integrin alphaM
 CC subunit precursor protein of the invention.
 XX
 XX Sequence 1152 AA;
 Query Match 99.58; Score 5846.5; DB 8; Length 1152;
 Best Local Similarity 99.28; Pred. No. 0;
 Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;
 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 60
 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 76
 61 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK 120
 77 RLQVPVEAVNMSLGLSLAATSPQLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK 136
 121 FPEALRGCPQEDSDIAPLVDSGSIIPHDFFRAKEFISTVMEQLKSKTILFSLMQVSEEF 180
 137 FPEALRGCPQEDSDIAPLVDSGSIIPHDFFRMKEFISTVMEQLKSKTILFSLMQVSEEF 196
 181 RIHFTKFEQNNPNRSLIKPIITQLGRTHTATGIRKVVRELFNITNGARKNAFKILILI 240
 197 RIHFTKFEQNNPNRSLIKPIITQLGRTHTATGIRKVVRELFNITNGARKNAFKILIVI 256
 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300
 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 316
 301 NFALKTIQNLREKIFAISGTQTGSSSSFEHMSQEGFSAAITSNGPLLLSTVGSDYDAG 360
 317 NFALKTIQNLREKIFAISGTQTGSSSSFEHMSQEGFSAAITSNGPLLLSTVGSDYDAG 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGVAMFR 420
 DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGVAMFR 436
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQGVSVCP 480
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGQGVSVCP 496
 QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEEEDNRGAVYLF 540
 DB 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEEEDNRGAVYLF 555
 QY 541 HGTSGSGISPSHSORIASKLSRLOVFGOSLGGQDLTMDGLVDLTGAGCHVILLRSQ 600
 DB 556 HGTSGSGISPSHSORIASKLSRLOVFGOSLGGQDLTMDGLVDLTGAGCHVILLRSQ 615
 QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVOKSTRDLRREGIOQSVVT 660
 DB 616 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVOKSTRDLRREGIOQSVVT 675
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
 DB 676 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
 QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSTIFSMSLDCLVVG 780
 DB 736 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSTIFSMSLDCLVVG 795
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSQRWRLACESASTEV 840
 DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSQRWRLACESASTEV 855
 QY 841 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
 DB 856 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHOYOVNLSGORSIPISVFLVPV 960
 DB 916 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHOYOVNLSGORSIPISVFLVPV 975
 QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1020
 DB 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1035
 QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFFVRSOTET 1080
 DB 1036 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQGAFFVRSOTET 1095
 QY 1081 KVBPPFEVNPPLIIVGSSVGLLLALLITAAALYKLGFFKQYKDMMSGEGPGGAEPQ 1137
 DB 1096 KVBPPFEVNPPLIIVGSSVGLLLALLITAAALYKLGFFKQYKDMMSGEGPGGAEPQ 1152
 RESULT 9
 ADP12435
 ID ADP12435 standard; protein; 1152 AA.
 XX
 AC ADP12435;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Protein encoded by mRNA of the invention #45.
 XX
 KW transplant rejection; immune system; rheumatoid arthritis; lupus;
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN W02004042346-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 24-APR-2003; 2003WO-US012946.

XX 24-APR-2002; 2002US-001311831.
PR 20-DEC-2002; 2002US-00325899.
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX WPI; 2004-400724/37.
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX Claim 65; SEQ ID NO 2444; 1762pp; English.
XX The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX Sequence 1152 AA;
SQ
Query Match 99.4%; Score 5839.5; DB 8; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;
QY 1 FNLDTENAMTQENARGFQGSVVOLOGSRVVVVGAPQETVAANORGLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTQENARGFQGSVVOLOGSRVVVVGAPQETVAANORGLYQCDYSTGSCPEI 76
QY 61 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQQPQK 120
DB 77 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFGLNLRQQPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTLFSLMOYSBEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKSKTLFSLMOYSBEF 196
QY 181 RIHFTFKEPQNNPRSLIKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILILJ 240
DB 197 RIHFTFKEPQNNPRSLIKPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSVTGSDYDAG 360
DB 317 NFEALKTIONLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSVTGSDYDAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILNRVQSLVLGAPYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILNRVQSLVLGAPYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAFYFASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQVSCPL 480
DB 437 QNTGMWESNANVKGTQIGAFYFASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQVSCPL 496
QY 481 PRGQARWQCDVLYGEOGQPKWRFGAALTIVLDVNGDKLTDVAIGAPGEENRGAIVLF 540
DB 497 PRG-RARWQCDVLYGEOGQPKWRFGAALTIVLDVNGDKLTDVAIGAPGEENRGAIVLF 555

QY 541 HGTSGGISPSHSORIASGLSPRIQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 600
DB 556 HGTSGGISPSHSORIASGLSPRIQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLRSQ 615
QY 601 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
DB 616 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETHKSTRRTQVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 720
DB 676 YDLALDSGRPHSRVAFNETHKSTRRTQVGLGTOTCETLKLQLPNCIEDPSPVILRLNF 735
QY 721 SLVGTPLSAFGNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 780
DB 736 SLVGTPLSAFGNLRPVLAEADQRLFTALFPEKKNCGNDNICODDLSITFSFMSLDCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 840
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 855
QY 841 SGALKSTSCSINHIFIPENSSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHIFIPENSSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHYQVSNLQORSLPISLVFLVPV 960
DB 916 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHYQVSNLQORSLPISLVFLVPV 975
QY 961 RLNQTIVWDRPQVTFSENLSTCHTKERLPKSHDFLAELRKAPVNVVNCIAVCQRIQCIP 1020
DB 976 RLNQTIVWDRPQVTFSENLSTCHTKERLPKSHDFLAELRKAPVNVVNCIAVCQRIQCIP 1035
QY 1021 PFGQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTE 1080
DB 1036 PFGQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTE 1095
QY 1081 KVEPFEVNPPLPLTVGSSVGGLLIALLITAAALYKLGFFKQYKDMWSEGGPPGAPPQ 1137
DB 1096 KVEPFEVNPPLPLTVGSSVGGLLIALLITAAALYKLGFFKQYKDMWSEGGPPGAPPQ 1152
RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX ADP44061;
XX 09-SEP-2004 (first entry)
XX Human CD11C protein SEQ ID NO:14.
XX HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX Homo sapiens.
XX WO2004053094-A2.
XX 24-JUN-2004.
XX 08-DEC-2003; 2003WO-US039208.
XX 06-DEC-2002; 2002US-0431522P.
XX (PPDP-) PPD DEV LP.
XX Dunn SJ, Holzmayer TA;
XX WPI; 2004-480928/45.
DR N-PSDB; ADP44060.
XX Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

Query Match				59.0%; Score 3469; DB 8; Length 1163;
Best Local Similarity				61.0%; Pred. No. 1.5e-280;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;				
QY	1	ENLDTENAMTFOENARFGQSVVLOGSRVVGAPQEIIVANQORGLYQCDYSTGSCPEI	60	
Db	20	FNLDTEELAFRVDVSAGFSDVQVANSVVVGAPOKITAANQOTGLYQCGYSTGACEPI	79	
QY	61	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQCSTENTYVKGICFLFGSNLQPOQK	120	
Db	80	GLQVPEAVNMSLGLSLAATTPPOLLACGPTVHQCSTENTYVKGICFLFGSNLQPOQK	137	
QY	121	FPFALRGCPQEDSDIAFLVDGSGSIIPHDPRAKEFISTVMEQLKKSKTLFSLMOYSBEF	180	
Db	138	LPVSRQECPRQDQVFLIDGSGSISSRNFAVMNFVRAVISQFQRPSTQFSLMQFSNKF	197	
QY	181	RHTFTKFEQNNPNSRLKPTQLGRTHATGKVKVRELFNITNGARKNAFKILILI	240	
Db	198	QHTFTFEBFRSTNSPLSLASVHQLQGFYTTATAIQNVVHRLFHASYGARRDATKILIVI	257	
QY	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNNTVASKPPRDHVQIN	300	
Db	258	TDGKKEGSLDYKQVWPADDAAGIIRYALGVGLAFQNRNSMKELNDIASKPSQEHIFVE	317	
QY	301	NPEALKTTONQUREKIFAIEGTQTCSSSFEHMSQEGFSAAITNGPILLSVTGSDYDAG	360	
Db	318	DFDALKDIONQLKEKIFAIEGTQTCSSSFEHMSQEGFSAAITNGPILLSVTGSDYDAG	377	
QY	361	GVELYTSKESTFTNMTRVDSMDNDAYLYGAAIILNRVQSLVLGAPRYQHIGLVAMFR	420	
Db	378	GAFLYPPNMSPTFINMSQENVMDRSDYLYGSTELALMKGVQSLVLGAPRYQHTGKAVFT	437	
QY	421	QNTGWESNANVKGTOIGAYGASCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSVCP	480	
Db	438	QVSRQWRMKAETVGTQISYFGASCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSVCP	497	
QY	481	PRGQRQCCDAVLYGEGQCPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVLYF	540	
Db	498	PRGWR-RWCCDAVLYGEGQCPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNKGAVLYF	556	
QY	541	HGTSGSGTSPSHSQRISGLSPRLQVPGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ	600	
Db	557	HGVLGPGSISPSHSQRISGLSPRLQVPGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ	616	
QY	601	PVLKVKALMEPNPREVARNPECDQVVKGEAGEVRVCLHVQKSTRDLREGIOISVVT	660	
Db	617	PVLWVGVSQMFIPAEIPRSAFECEQVVEQTLVQSNICLYDKRKNLGSRLQSSVT	676	
QY	661	YDLALDSGRHSRAVFNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSPVLRINF	720	
Db	677	LDLALDPGLSPRATFQETKNRSLSRVRLGLKAHCENFLLPSVCVEDSVTPITRLNF	736	
QY	721	SLVGTPLSAFNLRLPVLAEADQRLFTALFPPEKNCNDNTICQDLSITFTFMSLDCLVWG	780	
Db	737	TLVGKPLLAFLNRUPMLAALQRYFTASLPPEKNCNADHICQDNLGIFSFPGLKSLVG	796	
QY	781	GPREFNVTVVRNDGDSYKQTVTFPPDLDSYRKVSTLQNRQSRWSRLACESASSTEV	840	
Db	797	SNLELNAEVMVNDGDSYGTITFSPAGLSYRYVAEGQKQGLRSLHLTCDSPAPVG--	854	
QY	841	SGALKSTSCSINHPIFPENSEVFTNITFDVDSKASLGNKLLKANVTSENNMPRTKTEF	900	
Db	855	SGQWTSTSCRINHPIFRGGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTSKTF	914	
QY	901	QLELPVKYAVVYVTSHTGVTKYLNFAS-ENTSRVMOHQVQVSNLQCRSLPTSLVFLVP	959	
Db	915	QLELPVKYAVVYVTSHTGVTKYLNFAS-ENTSRVMOHQVQVSNLQCRSLPTSLVFLVP	974	
QY	960	VLNQTIVWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSAVQCRQCDI	1019	
Db	975	VELNQAEVMDVESHVQPNFLRCSSEKIAPPASDFLAHQKPNFVLDSCSIAGCLRFCVD	1034	
QY	1020	PFQIOBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGGAFVRSQTE	1079	

Db	1035	PSFSVQEEELDFTLKGNLSFGWVRQILQKKVSVVSAEITFTDTSVYSQLPQGEAFMEAQTT	1094	
QY	1080	TKVEPEPEVENPLIVGVSSVGLLLALITAAALYKLGFFKROVKDMMSE	1128	
Db	1095	TVLEKYKVHNPTPLIVGVSSIGLLJLALITAVLYKVGFFKROVKEMWEE	1143	
RESULT 13				
ID	AAR07120			
XX	AC	AAR07120 standard; protein; 1163 AA.		
XX	AC	AAR07120;		
XX	DT	25-MAR-2003 (revised)		
DT	05-FEB-1991	(first entry)		
XX	DE	p150.95 alpha subunit encoded by clone lambda dx47.		
XX	KW	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia; rhinovirus.		
KW	OS	Synthetic.		
XX	Key	Location/Qualifiers		
PH	Region	1..19		
FT	Region	/label= signal peptide		
FT	Region	20..44		
FT	Modified-site	/label= N-terminus		
FT	Modified-site	61..63		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	89..91		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	385..387		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	392..394		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	697..699		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	735..737		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	899..901		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	904..906		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	939..941		
FT	Modified-site	/label= glycosylation site		
FT	Modified-site	1050..1052		
FT	Modified-site	/label= glycosylation site		
FT	Domain	1108..1133		
FT	Domain	/label= transmembrane		
XX	WO9010646-A.			
PN	XX	20-SEP-1990.		
XX	PF	09-MAR-1990; 90WO-US001257.		
XX	XX	09-MAR-1990; 90WO-US001257.		
XX	PA	(DAND) DANA FARBER CANCER INST INC.		
XX	PI	Corbi AA, Springer TA;		
XX	XX	WPI; 1990-304985/40.		
DR	DR	N-PSDB; AAQ06068.		
XX	PT	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-unit of p150.95 cell surface adhesion receptor, opt. together with a beta chain of CD-18 family.		
XX	PS	Disclosure; Fig 3; 59pp; English.		

4

Qy	1080	TKEPPEPVNPPLPIVSSVGILLALITAAALYKLGFFRKQYKDMMSE	1128
Dd	1095	TVLEKKYVNHPDPIIVGSSIGGLLALITAVLYKGVFFRKQYKEMWEE	1143
RESULT 15			
AAB07361	ID	AAB07361 standard; protein; 1163 AA.	
XX	AC	AAB07361;	
XX	DT	17-JAN-2001 (first entry)	
XX	DE	Human CD11c protein sequence.	
XX	KW	Human; macrophage infiltration inhibition; alpha d integrin;	
XX	KW	leukocyte integrin; Leu-CAM; leukointegrin; immune response;	
XX	KW	inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;	
XX	KW	atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;	
XX	KW	lung inflammation; acute respiratory distress syndrome; Crohn's disease;	
XX	KW	rheumatoid arthritis; central nervous system injury; CD11c.	
OS	HO	Homo sapiens.	
XX	OS		
XX	WO	2000029446-Al.	
PN	XX		
PD	XX	25-MAY-2000.	
PF	XX	16-NOV-1999; 99WO-US027139.	
PR	XX	16-NOV-1998; 98US-00193043.	
PR	XX	08-JUL-1999; 99US-00350259.	
PA	XX	(ICOS-) ICOS CORP.	
PI	XX	Gallatin MW, Van Der Vieren M;	
PT	XX	WPI; 2000-387751/33.	
PT	XX	Use of novel anti-alpha integrin d monoclonal antibodies to inhibit	
PT	XX	macrophage infiltration and reduce inflammation at central nervous system	
PT	XX	injury sites.	
PS	XX	Example 5; Fig 1; 270pp; English.	
CC	CC	Integrins are a class of membrane-associated molecules that participate	
CC	CC	in cellular adhesion. Integrins are made up of an alpha subunit and a	
CC	CC	beta subunit. One class of human integrins are restricted to expression	
CC	CC	in white blood cells and have a common beta2 subunit: the leukocyte	
CC	CC	integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins	
CC	CC	have an important role in immune and inflammatory responses. The present	
CC	CC	protein sequence is the human integrin alpha subunit CD11c. This sequence	
CC	CC	was used in an alignment to identify a novel beta2 integrin alpha	
CC	CC	subunit: alpha d (AAA60014 and AAB07359). The present sequence has	
CC	CC	approximately 66% identity to the protein sequence of alpha.d. The	
CC	CC	Alpha.d gene and protein may be useful in therapy for diseases linked to	
CC	CC	aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple	
CC	CC	sclerosis, asthma, psoriasis, lung inflammation, acute respiratory	
CC	CC	distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency	
CC	CC	(LAP). In addition, anti-alpha d monoclonal antibodies may be used in the	
CC	CC	inhibition of macrophage infiltration at the site of a central nervous	
CC	CC	system injury. The monoclonal antibodies can also be used to detect and	
CC	CC	diagnose Crohn's disease	
XX	XX		
XSQ	XX	Sequence 1163 AA;	

Query Match 58.6%; Score 3446; DB 3; Length 1163;
Best Local Similarity: 60.9%

Query Match 58.6%; Score 3446; DB 3; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.3e-278;

Matches		686;	Conservative	147;	Mismatches	290;	Indels	6;	Gaps	4;
Qy	1	FNLDTENAMTFOENARFGQSVVQIQSGSVVVGAPQOEIVAAOQSGLYQCDYSTGSCPEI 60								
Db	20	FNLDTTEELTAFRVDSAGFGDSVVOYANSVVVVGAPQKIIAANQIGGLYQCGYSTGACEPI 79								
Qy	61	RLOVPVVEAVNMSLGLSLAATTSBPQIACGPTVHTQTCSENTYVVKGLCFLFGSNLRQQPQK 120								
Db	80	GLOVPEPEAVNMSLGLSLASTTSPSQLACGPTVHHECGRNMYLTGLCFLLGPT--QLTOR 137								
Qy	121	PPEALRGCPQBDSDIAFLVDGSGSIIPHDFFRAKEFIITVMEQLKKSKTLFSLMQYSEEF 180								
Db	138	LPVSRQECPRQBDIVFLIDGSGSISSRNFAITMNFVRAVISQFORPSTQFSLMQFSNKF 197								
Qy	181	RIHFTFKFQNNPNRSLKIPITQLLGRTHATGIRKVVRELPNITNGARKNAFKILILI 240								
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDAIKILIVI 257								
Qy	241	TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAPFRSEKSRQELNVTASKPPRDHVFQIN 300								
Db	258	TDGKKEGDSLDYKQVIPWADAAGIIRYALGVGLAFQNRNSWKELNDIAKPSOEHI FKVE 317								
Qy	301	NFEALKTIQNLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPELLSTVSGYDWAG 360								
Db	318	DFDALKDIONQLKEKIFAIEGTETISSSFLEMAQEGFSAVFTPDGPVLGAVGFTWSG 377								
Qy	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILRNEVQSLVIGAPRYOHIGLVAMPR 420								
Db	378	GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVIGAPRYOHIGKAVIFI 437								
Qy	421	QNTGWNESNANVKGTOIGAYFGASLCSDVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480								
Db	438	QVSRQWRMKAEEVIGTOIGSYFGASLCSDVDVDGSTDVLIGAPHYYEQTRGGQVSCPL 497								
Qy	481	PRGQBARWOCDAVLXGEOQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540								
Db	498	PRGWR--RWCDAVLXGEOQHPWGRFGAALTIVLGDVNGDKLTDVVI GAPGEENRGAVILF 556								
Qy	541	HGTSSGSI SPSSHQSRIAGSKLPRLOYFGOSISGGDLTMDGLVDLTVCAQGHVLLRSQ 600								
Db	557	HGVLGFSISPSHQSRIAGSKLPRLOYFGQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616								
Qy	601	PVLRVKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGCQIQSVVT 660								
Db	617	PVLWGVSMQFTPAEIPRAFAECREQVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676								
Qy	661	YDLALDSGRPHSAFNETKNSTRQTQVLGTLQTCTETLKLQLPNCIEDPVPVLRINF 720								
Db	677	LDLALAPGRLSPRAI FOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVIPIILRLNF 736								
Qy	721	SLVGTPLSAFNLRPVLAEDAQLFTALPFEKNCGNDNICODDLSITFSEMSLDCLVVG 780								
Db	737	TLVGKPELLAPNLRPLAALAQRYFTASLPFEKNCAGDHCQDNLGISFSFPGKSLLVG 796								
Qy	781	GPREFNVTVVRNDGDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840								
Db	797	SNLELNAEVMVWMDGEDSTGTTITFSPHAGLSRYVYVAGQKQQLRSLHLTC--CSAPVG 854								
Qy	841	SGALKSTSCSINHPIFENSEVTNITFDVDSKASIGNKLLKXANTVSENMPRTNKTFF 900								
Db	855	SQGTWSTSCRINHLIFRGGAQITFLATFDVSPRAVGLDRLLLIANTVSSENNI PRTSKITF 914								
Qy	901	QLELPVYAVYVTVTSHGVSTKYLNFTAS--ENTSRVMQHOYQVSNLQORSLPISLFLVP 959								
Db	915	QLELPVYAVYIVVSSHEQFTKYLNFSSESEKESHVAMHRYQNNLQORDLPVSIWFVP 974								
Qy	960	VRLNQTVIMDRPQVTFSENLSCTHTKRLPSHDFLAELRKAPVNVNCSTIAVCQRIQCDI 1019								
Db	975	VELNQEAVMVDVEVSHQNFSLSCSEKTIAPPASDFLAHIQKNPVLDCSIAGCLRFRCDV 1034								
Qy	1020	PFFGIQDEFNATLKNLSFDWYIKTSHNHLIIIVSTAEILLFNDSVFILLPGQAFVRSOTE 1079								
Db	1035	PSFSVQBELDFTLKNLSFGWVRQILQKKVSVSVABIIIFDTSVYSQLPQGAQPMRAQTI 1094								

Qy 1080 TKVEPFVFPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMMS 1128
Db 1095 TVLEKYKVHNPPIPIVGVSSIGLLLLALITAVLYKVGFFKRYKEMMEE 1143

Search completed: November 9, 2004, 12:08:24
Job time : 108.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-5
Perfect score: 5876
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*

7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*

9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*

10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*

11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*

12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*

13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*

14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*

15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*

16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pcp.*

17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*

18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pcp.*

19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*

20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5876	100.0	1137	10	US-09-902-481A-5
2	5862	99.8	1153	9	US-09-350-259-3
3	5862	99.8	1153	10	US-09-902-481A-1
4	5862	99.8	1153	10	US-09-891-943-3
5	5862	99.8	1153	14	US-10-144-259-30
6	5862	99.8	1153	14	US-10-207-655-176
7	5855	99.6	1137	10	US-09-902-481A-6
8	5851	99.6	1137	10	US-09-902-481A-4
9	5846.5	99.5	1152	9	US-09-945-265-4
10	5836	99.3	1137	10	US-09-902-481A-3
11	3469	59.0	1163	14	US-10-116-275-204
12	3446	58.6	1163	9	US-09-350-259-4
13	3446	1163	10	US-09-891-943-4	

14	3411	58.0	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3411	58.0	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3395.5	57.8	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3395.5	57.8	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3232.5	55.0	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3232.5	55.0	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3222	54.8	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3222	54.8	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3211.5	54.7	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3211.5	54.7	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3201	54.5	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3201	54.5	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1557.5	26.5	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1553.5	26.4	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1530	26.0	1223	16	US-10-408-765A-295	Sequence 295, App
30	1359.5	23.1	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1153	19.6	1179	14	US-10-177-550-2	Sequence 2, Appli
34	1153	19.6	1179	14	US-10-173-551-2	Sequence 101, App
35	1151.5	19.6	413	9	US-09-350-259-101	Sequence 101, App
36	1151.5	19.6	413	10	US-09-891-943-101	Sequence 103, App
37	1103.5	18.8	1151	10	US-09-884-130-103	Sequence 103, App
38	1103.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
39	1103.5	18.8	1179	10	US-09-918-715-250	Sequence 250, App
40	1103.5	18.8	1179	17	US-10-474-794-250	Sequence 250, App
41	1101	18.7	1188	14	US-10-291-265-810	Sequence 810, App
42	1097	18.7	1188	14	US-10-291-265-338	Sequence 338, App
43	1093.5	18.6	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1093.5	18.6	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1093.5	18.6	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-902-481A-5

; Sequence 5, Application US/09902481A

; Publication No. US20030054440A1

; GENERAL INFORMATION:

; APPLICANT: Springer, Timothy

; APPLICANT: Shimooka, Motomu

; APPLICANT: Shifman, Julia

; APPLICANT: Mayo, Stephen

; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY

; FILE REFERENCE: A-70586-1/RFT/RMS/RMK

; CURRENT APPLICATION NUMBER: US/09/902.481A

; CURRENT FILING DATE: 2001-07-09

; PRIOR APPLICATION NUMBER: US 60/216,600

; PRIOR FILING DATE: 2000-07-07

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5

; LENGTH: 1137

; TYPE: PRT

; ORGANISM: Artificial sequence

; FEATURE:

; OTHER INFORMATION: synthetic

US-09-902-481A-5

Query Match	100.0%;	Score 5876;	DB 10;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;	
QY	1	FNLDTENAMTFQENARGFGQSVVOLGSRVVGAPQEIIVANORGSLYQCDYTGSCPEI	60	
Db	1	FNLDTENAMTFQENARGFGQSVVOLGSRVVGAPQEIIVANORGSLYQCDYTGSCPEI	60	
QY	61	RLQVPVEAVNMISGLSLAATTSPQQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	120	
Db	61	RLQVPVEAVNMISGLSLAATTSPQQLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK	120	

121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQKKSITLFSLMQYSEEF 180
121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQKKSITLFSLMQYSEEF 180
181 RIHFTFEKFPQNNPNSLIKPIITQLLGRTHATGIRKVVRELFNTNGARKNAFKILLI 240
181 RIHFTFEKFPQNNPNSLIKPIITQLLGRTHATGIRKVVRELFNTNGARKNAFKILLI 240
241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVFOIN 300
241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVFOIN 300
301 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
301 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
421 QNTGWMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTRGGQVSVCP 480
421 QNTGWMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTRGGQVSVCP 480
481 PRGQARWQCDVAVLYGEGQPGWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
481 PRGQARWQCDVAVLYGEGQPGWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGDLTMDGLVLTGVAQGHVLLRSQ 600
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGDLTMDGLVLTGVAQGHVLLRSQ 600
601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
661 YDLALDSGRPHRAVFNETHKSTRQTQVGLGTQTCETLKLQIPNCIEDPVSPIVLRNF 720
661 YDLALDSGRPHRAVFNETHKSTRQTQVGLGTQTCETLKLQIPNCIEDPVSPIVLRNF 720
721 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFEKNCNDNICODDLSITFSPMSLDCLVVG 780
721 SLVGTPLSAFGLNRPVLAEDAORLFTALPPFEKNCNDNICODDLSITFSPMSLDCLVVG 780
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVSTLQNRQSRWSRLACESASSTEV 840
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVSTLQNRQSRWSRLACESASSTEV 840
841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTEF 900
841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTEF 900
901 QLELPVKYAVMVVTVSHGVSTKYNFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
901 QLELPVKYAVMVVTVSHGVSTKYNFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1020
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1020
1021 FPGIOEENATLKNLSFDWYIKTSHNLLIYSTABILFNDVSFTLLPGOGAFVRSQTET 1080
1021 FPGIOEENATLKNLSFDWYIKTSHNLLIYSTABILFNDVSFTLLPGOGAFVRSQTET 1080
1081 KVEPPEFVNPPLPIVGVSSVGGILLIALLIITAALYKLGFFKRYQKDMMSSEGGPPGAE 1137
1081 KVEPPEFVNPPLPIVGVSSVGGILLIALLIITAALYKLGFFKRYQKDMMSSEGGPPGAE 1137

RESULT 2
US-09-350-259-3
; Sequence 3, Application US/09350259

; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; EARLIER FILING DATE: 1999-07-08
; EARLIER FILING DATE: 1998-11-16
; EARLIER FILING DATE: 1993-12-23
; EARLIER FILING DATE: 1993-12-23
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-08-05
; EARLIER FILING DATE: 1994-12-21
; EARLIER FILING DATE: 1994-12-21
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.8%; Score 5862; DB 9; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTQENARGFGQSVVQLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCBPI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLSNLRQPOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLSNLRQPOK 136
QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQKKSITLFSLMQYSEEF 180
DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQKKSITLFSLMQYSEEF 196
QY 181 RIHFTFEKFPQNNPNSLIKPIITQLLGRTHATGIRKVVRELFNTNGARKNAFKILLI 240
DB 197 RIHFTFEKFPQNNPNSLIKPIITQLLGRTHATGIRKVVRELFNTNGARKNAFKILLI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVFOIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVFOIN 316
QY 301 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGWMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTRGGQVSVCP 480
DB 437 QNTGWMESNANVKGTOIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEGQPGWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEGQPGWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGDLTMDGLVLTGVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQGSLSGGDLTMDGLVLTGVAQGHVLLRSQ 616
QY 601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676

```

QY 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLGRSLPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLGRSLPLSLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWIKTSNNHLLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWIKTSNNHLLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096
QY 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

```

RESULT 3

```

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimada, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)-()
; OTHER INFORMATION:
US-09-902-481A-1

```

```

Query Match 99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 FNLDTENAMTFQENARFGQSVWLQGSRRVVGAPQBEIVANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVWLQGSRRVVGAPQBEIVANQRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLACGTVHTQTCSENTYVVKGLCFPLFGSNLRQOPQK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLACGTVHTQTCSENTYVVKGLCFPLFGSNLRQOPQK 136

```

RESULT 4

```

US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US2003007728A1

```

```

QY 121 FPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFIITVMQOLKSKTLFSLMOYSEBF 180
Db 137 FPEARLGCPOEDSDIAFLVDGSGSIIPHDFFRAKEFIITVMQOLKSKTLFSLMOYSEBF 196
QY 181 RHITFFKFXFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTINGAKNAFKILILI 240
Db 197 RHITFFKFXFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTINGAKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNVTASKPPDRHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNVTASKPPDRHVFQIN 316
QY 301 NFEALKTQNLREKIFAIEGTQGTGSSSFHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTQNLREKIFAIEGTQGTGSSSFHEMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYOHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
QY 481 PRGORARQCCDAVLYGEOGQPWGRFGAALTIVLGDWNGDKLTDVAIGAPCEEDNRGAVYLF 540
Db 497 PRGORARQCCDAVLYGEOGQPWGRFGAALTIVLGDWNGDKLTDVAIGAPCEEDNRGAVYLF 556
QY 541 HGTSGSGTSPSHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
Db 557 HGTSGSGTSPSHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
QY 601 PVLVKAIMFPNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQISVVT 660
Db 617 PVLVKAIMFPNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQISVVT 676
QY 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLGRSLPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLGRSLPLSLVFLVPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVCSIAVCQRIQCDDIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWIKTSNNHLLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWIKTSNNHLLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096
QY 1081 KVEPFEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
Db 1097 KVEPFEVNPPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

```

```

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-05-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-3

Query Match          99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEQLKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEQLKSKTLFSLMQYSEEF 196
QY 181 RHFTFKFQNNPNPSLKIPTQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILLI 240
DB 197 RHFTFKFQNNPNPSLKIPTQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILLI 256
QY 241 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVQFN 300
DB 257 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVGDAFSEKSRQELNITVASKPRDHVQFN 316
QY 301 NFEALXTIONQLREKIFAIEGTQSGSSFEHMSQEGFSAATISNGPLLLSTVGSYDMAG 360
DB 317 NFEALXTIONQLREKIFAIEGTQSGSSFEHMSQEGFSAATISNGPLLLSTVGSYDMAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLGAPRYQHIGLVAMER 420
DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLGAPRYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL 496
QY 481 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540
DB 497 PRQORARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVKQSTDRLEQIOQSVVT 660
DB 617 PVLRVKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVKQSTDRLEQIOQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
```

```

DB 677 YDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKKGNDNICODDLSTIFFSFWSLDCLVVG 780
DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKKGNDNICODDLSTIFFSFWSLDCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSYRTQVTFPPFPDLDSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVYVTVTSHGVSTKYLNFNTASENTRVWQHQQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVYVTVTSHGVSTKYLNFNTASENTRVWQHQQYQVSNLQSRSLPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 PFGIOEBFNATLKGNSLFDWIKTSHNHLIIVSTAEILLFNDVFTLLPCOGAFVRSOTET 1080
DB 1037 PFGIOEBFNATLKGNSLFDWIKTSHNHLIIVSTAEILLFNDVFTLLPCOGAFVRSOTET 1096
QY 1081 KVEPPEVNPPLIIVGSSVGGLLLALITAAALYKLGFPKQYKQKDMWSEGGPPGABPQ 1137
DB 1097 KVEPPEVNPPLIIVGSSVGGLLLALITAAALYKLGFPKQYKQKDMWSEGGPPGABPQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-144-259-30

Query Match          99.8%; Score 5862; DB 14; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEQLKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFIPTVMEQLKSKTLFSLMQYSEEF 196
QY 181 RHFTFKFQNNPNPSLKIPTQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILLI 240
DB 181 RHFTFKFQNNPNPSLKIPTQLLGRTHRTATGIRKVVRELFNITNGARKNAFKILLI 240
```

```
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVLGVGDARFSEKSRQELNVTASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVLGVGDARFSEKSRQELNVTASKPPRDHVFQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDWAQ 360
Db 317 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDWAQ 376
QY 361 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 617 PVLRVKALMEFNPVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSRRQTQVLGLTQTCETLKLQPNCIDBPVSPVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKNSRRQTQVLGLTQTCETLKLQPNCIDBPVSPVLRNLF 736
QY 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVTMVTSYHGVSTKYNFTASENTSRVMOHQVSNLQORSUPISLVFLVPV 960
Db 917 QLELPVKYAVTMVTSYHGVSTKYNFTASENTSRVMOHQVSNLQORSUPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP 1036
QY 1021 FFGIOEFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQ 1080
Db 1037 FFGIOEFNATLGNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQ 1096
QY 1081 KVEPPEVPNPLIVGSSVCGHLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
Db 1097 KVEPPEVPNPLIVGSSVCGHLLALITAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153
```

RESULT 6

```
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
```

```
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176
```

Query Match

```
Best Local Similarity 99.8%; Score 5862; DB 14; Length 1153;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAANQORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAANQORGLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGPTVHTQTCSENTYVVKGLCFGLPSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGPTVHTQTCSENTYVVKGLCFGLPSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSII PHDFERRAKEFISTVMEOLKSKTLFSLMOYSEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSII PHDFERRAKEFISTVMEOLKSKTLFSLMOYSEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVLGVGDARFSEKSRQELNVTASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVLGVGDARFSEKSRQELNVTASKPPRDHVFQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDWAQ 360
Db 317 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDWAQ 376
QY 361 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILLRNRVQSLVIGAPYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTRGQVSVCP 496
QY 481 PRGQARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 617 PVLRVKALMEFNPVARNVFECDNVVKGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSRRQTQVLGLTQTCETLKLQPNCIDBPVSPVLRNLF 720
Db 677 YDLALDSGRPHSRVAVNETKNSRRQTQVLGLTQTCETLKLQPNCIDBPVSPVLRNLF 736
QY 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDRQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
```

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVELVPV 960
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVELVPV 976
QY 961 RLNTQVWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020
DB 977 RLNTQVWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1036
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAAILFNDVSFTLLPGQAFVRSQTET 1080
DB 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAAILFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVEPEVFNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKQVNMSEGGPPGAEPQ 1137
DB 1097 KVEPEVFNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKQVNMSEGGPPGAEPQ 1153

RESULT 7
US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match 99.6%; Score 5855; DB 10; Length 1137;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
DB 1 ENLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
QY 61 RLQVPVEAVNMSLGLSAAATSPQLLACGPTVHTQCSNTYVYKGLCFGLFSGNLROQPK 120
DB 61 RLQVPVEAVNMSLGLSAAATSPQLLACGPTVHTQCSNTYVYKGLCFGLFSGNLROQPK 120
QY 121 FPEALGCCPEDSDIAPLVNDSGSIIPHDERRAKEFTSTVMEQKKSKTLFSLMQSSEEF 180
DB 121 FPEALGCCPEDSDIAPLVNDSGSIIPHDERRAKEFTSTVMEQKKSKTLFSLMQSSEEF 180
QY 181 RIHFTKEFQNNPNSLKIPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 181 RIHFTKEFQNNPNSLKIPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKRSQELNIVASKEPRDHVQFN 300
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKRSQELNIVASKEPRDHVQFN 300
QY 301 NFEALKTIQNLREKIFAIRGTGTGSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG 360
DB 301 NFEALKTIQNLREKIFAIRGTGTGSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFFINMTFVDSMDNDAYLGVAIAIILNRRVQSLVLGAPRYOHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFFINMTFVDSMDNDAYLGVAIAIILNRRVQSLVLGAPRYOHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480
DB 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480
QY 481 PRGQARWQCDVAVLYGSGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNKGAVYLF 540
DB 481 PRGQARWQCDVAVLYGSGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNKGAVYLF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSISGGQDITMDGLVDLTVGAOCHVLLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSISGGQDITMDGLVDLTVGAOCHVLLLRSQ 600
QY 601 PVLRVKALMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKALMEFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVAFNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKSTRTQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720
QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDCGDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWRLACESASSTEV 840
QY 841 SGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVELVPV 960
DB 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVELVPV 960
QY 961 RLNTQVWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020
DB 961 RLNTQVWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAAILFNDVSFTLLPGQAFVRSQTET 1080
DB 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIIIVSTAAILFNDVSFTLLPGQAFVRSQTET 1080
QY 1081 KVEPEVFNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKQVNMSEGGPPGAEPQ 1137
DB 1081 KVEPEVFNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKRYQKQVNMSEGGPPGAEPQ 1137

RESULT 8
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic

US-09-902-481A-4

Query Match 99.6%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANGSLYQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120
DB 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEFTSTVMEOLKSKTFLFSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEFTSTVMEOLKSKTFLFSLMOYSEEF 180

QY 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240

QY 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300

QY 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480

QY 481 PRGQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

QY 541 HGTSGSGLSPHSQRIAGSKSLPRLOYFGSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGLSPHSQRIAGSKSLPRLOYFGSLSGQDITMDGLVDLTVGAQGHVLLRSQ 600

QY 601 PVLRVKALMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKALMEFPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPSPVILRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCLIEDPSPVILRLNF 720

QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQODDLSITFSPMSLDCLVVG 780
DB 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQODDLSITFSPMSLDCLVVG 780

QY 781 GPREFNVTVVRNDGDSYKTVTFPPFLDLSYKVKSTLQNRQSRWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGDSYKTVTFPPFLDLSYKVKSTLQNRQSRWRLACESASSTEV 840

QY 841 SGALKSTSCSINHPIFPENSEVFNTITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIFPENSEVFNTITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

QY 901 QLELPVKYAYVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPI SILVFLPV 960
DB 901 QLELPVKYAYVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPI SILVFLPV 960

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLP SHSDFLAELRKAPVNCVSIACQRIQCDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLP SHSDFLAELRKAPVNCVSIACQRIQCDIP 1020
```

```
QY 1021 FPGIQEFNATLKGNLISFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQGAFFRSOTET 1080
DB 1021 FPGIQEFNATLKGNLISFDWYIKTSHNHLIVSTAELFNDSVFTLLPGQGAFFRSOTET 1080

QY 1081 KVEPFEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKFOYKDMMEGPPGAEPPQ 1137
DB 1081 KVEPFEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKFOYKDMMEGPPGAEPPQ 1137
```

RESULT 9

```
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Iu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

Query Match 99.5%; Score 5846.5; DB 9; Length 1152;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEFTSTVMEOLKSKTFLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEFTSTVMEOLKSKTFLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 316

QY 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496

QY 481 PRGQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
```

541 HGTSGSGISPSHRSORIRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 600
556 HGTSGSGISPSHRSORIRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 615
601 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
616 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 675
661 YDLALDSGRPHSRVAFNETKSTRRTQTVLGTQTCETLKLQLPNCIEDPVSPIVRLNF 720
676 YDLALDSGRPHSRVAFNETKSTRRTQTVLGTQTCETLKLQLPNCIEDPVSPIVRLNF 735
721 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
736 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 795
781 GPREFNVTVVRNDGEDSVRTQVTFPPDLISYRKVSTLONORSQSWRLACESASSTEV 840
796 GPREFNVTVVRNDGEDSVRTQVTFPPDLISYRKVSTLONORSQSWRLACESASSTEV 855
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 915
901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960
916 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 975
961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
976 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035
1021 FFGIOEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFFVRSQTEP 1080
1036 FFGIOEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFFVRSQTEP 1095
1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137
1096 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1152

RESULT 10
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimooka, Motomu
; APPLICANT: Shifman, Julia
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.3%; Score 5836; DB 10; Length 1137;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVVGAPQEIIVANQORSLYQCDSYTGSCBPI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVVGAPQEIIVANQORSLYQCDSYTGSCBPI 60

QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACQPTVHQTCSNTYVKGCLFCLFSGNSLRQPOK 120
DB 61 RLQVPVEAVNMSLGLSLAATTPPQLLACQPTVHQTCSNTYVKGCLFCLFSGNSLRQPOK 120
121 FPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEFISTVMEOLKSKTFLSLMOYSBEF 180
DB 121 FPEALRGCPQSDSIAFLVDGSGSIIIPHDPRAKEFISTVMEOLKSKTFLSLMOYSBEF 180
181 RIHETFEFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAKILILI 240
DB 181 RIHETFEFQNNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAKILILI 240
241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNVTASKPRDHVFQIN 300
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELNVTASKPRDHVFQIN 300
301 NFEALKTQNLREKIPAIETGOTQSSSSFEHMSOEGFSAITNGPLLSVGVSDWAG 360
DB 301 NFEALKTQNLREKIPAIETGOTQSSSSFEHMSOEGFSAITNGPLLSVGVSDWAG 360
361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYEQTRGGQSVCP 480
DB 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLICAPHYEQTRGGQSVCP 480
481 PRGQARWOCDAVLGEQGPWRGGAALTVDLVGNDKLTDAIAGAPDEENRGAVYLF 540
DB 481 PRGQARWOCDAVLGEQGPWRGGAALTVDLVGNDKLTDAIAGAPDEENRGAVYLF 540
541 HGTSGSGISPSHRSORIRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 600
DB 541 HGTSGSGISPSHRSORIRAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGHVLRLRSQ 600
601 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
DB 601 PVLRVKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
661 YDLALDSGRPHSRVAFNETKSTRRTQTVLGTQTCETLKLQLPNCIEDPVSPIVRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKSTRRTQTVLGTQTCETLKLQLPNCIEDPVSPIVRLNF 720
721 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFNLRLPVLAEQAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
781 GPREFNVTVVRNDGEDSVRTQVTFPPDLISYRKVSTLONORSQSWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLISYRKVSTLONORSQSWRLACESASSTEV 840
841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
DB 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960
DB 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960
961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
1021 FFGIOEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFFVRSQTEP 1080
DB 1021 FFGIOEFNATLKGNSLDFWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFFVRSQTEP 1080
1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137
DB 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elna Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 59.0%; Score 3469; DB 14; Length 1163;
Best Local Similarity 61.0%; Pred. No. 2e-285;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGGQSVVQGGRRVVVGAPOEIVAAQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTENAMTFQENARGGQSVVQGGRRVVVGAPOEIVAAQORSLYQCDYSTGSCPEI 79
QY 61 RLQVPEAVNMISLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 120
Db 80 GLQVPEAVNMISLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSMOYSEF 180
Db 138 LPVSRQECRQEQDIFVLIDGSGSISSRNFMVFAVAVISQFQRPSTQSLMOFSNKF 197
QY 181 RIHETFEKQNNPRSLIKIPITQLLGHATGIRKVRVRELFNITNGARKNAFKILILI 240
Db 198 QTHETFEFRSTNPLSLASVHQLQGYTATATQNVVHLEFASYGARDATKILIVI 257
QY 241 TDGKFGDPLCYEDVPEADREGVIRYVIGVGDADFSEKSKROELNTVASKPRDHVQIN 300
Db 258 TDGKFGDPLCYEDVPEADREGVIRYVIGVGDADFSEKSKROELNTVASKPRDHVQIN 317
QY 301 NFEALKTQNLREKIFAEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 360
Db 318 DFDALKDIONLKEKIFAEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 377
QY 361 GVELYTSKEKSTFNMTVRDSDMDAYLGVAAAIILNRVQSLVGLAPRYQHI GLVAMFR 420
Db 378 GAFLYPPNMGPTFINMSQENVDSDYLGISTELALWKVQSLVGLAPRYQHTKAVIFT 437
QY 421 QNTGMWESNAVKGTQIGVAFGASLCSVDVDSNGSTDLVILGAPHYEQTGCGVSCPL 480
Db 438 QVSEQWEMKAEVTGTLGISTFGASLCSVDVDSNGSTDLVILGAPHYEQTGCGVSCPL 497
QY 481 PRGORARWCDAVLYGEGQGPWGRFGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 498 PRGWR-RWVCDVLYGEGQHPWGRFGAALTGLVDVNGDKLTDVVGAPGEEENRGAVYLF 556
QY 541 HGTSGSISPSHSQRIAGSKLSPQLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLGSPISPSHSQRIAGSKLSPQLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLVRKAIMFNFREAVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLVWVGSMQFIPAEIPRSFAFECREQVVSQTLVQSNICILYIDKRKNLIGSRDLQSSVT 676

RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vliet, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.6%; Score 3446; DB 9; Length 1163;
Best Local Similarity 60.8%; Pred. No. 1.8e-283;
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGGQSVVQGGRRVVVGAPOEIVAAQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTENAMTFQENARGGQSVVQGGRRVVVGAPOEIVAAQORSLYQCDYSTGSCPEI 79
QY 61 RLQVPEAVNMISLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 120
Db 80 GLQVPEAVNMISLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

QY 661 YDLALDSGRPHSRVAVENETKNTSTRQTVGLTQTCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 LDALDPGLSLPRATFQETKNSRLSVRVGLKAHCENFNLPLSPCEDSVTITLRLNF 736
QY 721 SLVGTPLSAFAGNLRLPVAEDAOQLFTALPFEKNCNNDNICODDLITITSFMSLDCILVVG 780
Db 737 TLVGPDLAFLAFLNRLPMLAALQRYFTASLPFEKNCNNDNICODDLITITSFMSLDCILVVG 796
QY 781 GPREFNVTVVRNDGSDSYRTQVTFPFLDLSYRKVSTLQNKORSQSRWRLACESASSTEV 840
Db 797 SNLELNAEVMVWVNDGSDSYRTQVTFPFLDLSYRKVSTLQNKORSQSRWRLACESASSTEV 854
QY 841 SGALKSTSCSINHPIPPENSEVFNITPVDVSKASIGNKLLKLLKANYTSENMPRTNKTBF 900
Db 855 SQGTWSTSCRIINHLIPRGAQITFLATFDVSPKAVGLDRLLLTANVSSNNPTSKITTF 914
QY 901 QLELPVKAVYVMTVSHGVSTKYANFTAS-ENTSRVMQHOYQVSNLQBSLPSLFLVLP 959
Db 915 QLELPVKAVYVMTVSHGVSTKYANFTAS-ENTSRVMQHOYQVSNLQBSLPSLFLVLP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVGSIACVORIODI 1019
Db 975 VELNQEAVMVDEVSHQPNQPSLRCSSEKIAPPASDFLAHQKVPVLDCSIAGCLPRCDV 1034
QY 1020 PFFGQIEFNATLKNLSFDMVYKTSNHLIVSTAEILFNDVSTFLPQCGAFVRSQTE 1079
Db 1035 PFSVQIEELDTLKNLSFDMVYKTSNHLIVSTAEILFNDVSTFLPQCGAFVRSQTE 1094
QY 1080 TKVPEPEVNPPLPLIVGSSVGLLILALITLALYKLGFFKQYKDMGE 1128
Db 1095 TVLEKYKHNTPLPLIVGSSVGLLILALITLALYKLGFFKQYKDMGE 1143

Db 677 LDALAPGLSPRAIFQETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVIILRLNF 736
QY 721 SLVGTPLSAFGLRPLVLAEDAORLFTALPFEKNCNGDNICQDLSITFSFMSLDCLVVG 780
Db 737 TLVGRFLAFLRNLPLMALAQRFTASLPFEKNCNGADHICQNLGISFSFGLKSLVVG 796
QY 781 GRPEFNVTVVRNDEGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWELACESASSTEV 840
Db 797 SNLELNAEVMVWVNDGSDSYGTTITFHPAGLSYRYVAEQKQGLRLHLTC--CSAPVG 854
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 SGTWSTSCRIHNLIFRGAQITFLATFDVSPKAVGLDRLLLTANVSSENNTPTSKTIF 914
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNF--ENTSRVMQHOYQVSNLQSRSLPISLFLVP 959
Db 915 QLELPVKYAVVMVTSHGVSSTKYLNF--ENTSRVMQHOYQVSNLQSRSLPISLFLVP 974
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFAELRKAPVNVCSIAVCQRIQCDI 1019
Db 975 VELNQEAVMMDVEVSHPNQPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIACCLPRCDV 1034
QY 1020 PFFGIQEEFNATLKNLFDWIKTSHNHLIVSTAEILFNDSTVTLPGQGAFFVRSQTE 1079
Db 1035 PSFSVQEEELDFTLKGLNLSFGWVRQILQKKVSVSVAEITFDTSVYSQLPQGEAFMRAQTI 1094
QY 1080 TKVEPFEPVNPPLVTVGSSVGLLILALITAAALYKLGFFKQYKDMWSE 1128
Db 1095 TVLEKYKHNPILVTVGSSVGLLILALITAAALYKLGFFKQYKDMWSE 1143

RESULT 14

US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-350-259-2
Query Match 58.0%; Score 3411; DB 9; Length 1161;
Best Local Similarity 59.4%; Pred. No. 1.7e-280;
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;
QY 1 ENLDTENAMTFQENARGFGSVVLOGSRVVGAPQEIIVAAVNRGLSYQCDYSTGSCPI 60
Db 17 FNLIDVEEPTIQBDAGFGQSVVQFGSRLVVGAPLEVAANQTRGLYDCAATGMCQPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHOTCSENYTKGCLFLFGSNLRQPOK 120
Db 77 PLHIREAVNMSLGLSLAATTPSPQLLACGPTVHOTCSENYTKGCLFLFGSNLRQPOK 135

QY 121 FPEALRGCOEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKSKTLFSLMQYSEEF 180
Db 136 VPDATEPECHQEMDIIVFLIDGSGSIDQDNFNQMGFVQAVMGQFEGTDTL.FALMQYSNLL 195
QY 181 RIHFTFKFQNNPNSRLKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 196 KIHFTFTQRTSFSQSLVDPIVQLKGLTFTATGILTVVTFHFKNGARKSAKILIVI 255
QY 241 TDGEKFGDPLGYEDYIPADREGVIRYVIGVDGDAFRSEKSRQELNITVASKPPDRHVFQIN 300
Db 256 TDGQYKXDFLEYSDDVIPAOKAGIIRYALGVGHAFQGTARQELNITISSAPQDHFVKVD 315
QY 301 NFEALKTTQNLQREKI FALGHTQTGSSSSFEHMSQEGFSAAITNGPLLLSTVGSYDWAG 360
Db 316 NFAALGSIQKQIQEXIYAVEGTQSRASSFQHEMSQEGFSTALTMDGLFLGAVGSFWSG 375
QY 361 GVFLYTSKEKSTFINWTRVDSMDNAYLGXAAIITLNRVQSLVILGAPRYQHILVAMPR 420
Db 376 GAFLYPPNMSPTFINNSQENVMDRSYLGYSTELALWKGVQNLVILGAPRYQHTGKAVIPT 435
QY 421 QNTGMWESNANVKGTIQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYETCTRGQGVSVCP 480
Db 436 QVSRQWRKKAETVGTIGSYFGASLCSVDVDSNGSTDLVLIGAPHYETCTRGQGVSVCP 495
QY 481 PRQORARWQCDVLYGEGQGWGPRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 496 PRQORVQWQCDVLYGEGQGWGPRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSISPSHSORIASGLSPRLQYFGALSQGGQDLTQDGLMDLAVGARGVQLLRLSL 615
Db 556 HGASEGISPSHSORIASGLSPRLQYFGALSQGGQDLTQDGLMDLAVGARGVQLLRLSL 615
QY 601 PVLKVAIMEFNPREVARNVFECDQVVKGEAGEVVRVCLHVQKSTDRDLREGOIQSVVT 660
Db 616 PVLKVGVMRPSPVEVAKAVVRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSVR 673
QY 661 YDLALDSGRPSRAVFNETKNSRTRQTVIGLGTCTETLKLQPNCTIEDPVSIVLRLNF 720
Db 674 YDLALDGRLLTSRAIFNETKNTPLTRRKTGLGTHCETKLLPDCEVEDVVSPIILHNF 733
QY 721 SLVGTPLSAFGLRPLVLAEDAORLFTALPFEKNCNGDNICQDLSITFSFMSLDCLVVG 780
Db 734 SLVREPISPNLRPLVAVGQDLTASLPFEKNCNGQDGLCEGDLGVTLSFSGLTITVG 793
QY 781 GPREFNVTVVRNDEGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWELACESASSTEV 840
Db 794 SSLELNVITVWVNDGSDSYGTTVSLYYPAGLSHRRVSGAOKQPHOSALRLACETV-PTED 852
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 853 EG-LRSSRCSVNHPIFHEGNSNGTFIVTFDVSYSKATLGDRLMLRASASSENKASSSKATF 911
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNF--TASENTRVMQHOYQVSNLQSRSLPISLFLVP 959
Db 912 QLELPVKYAVVMVTSHGVSSTKYLNF--TASENTRVMQHOYQVSNLQSRSLPISLFLVP 971
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFAELRKAPVNVCSIAVCQRIQCDI 1019
Db 972 VLLNGVAVMDVWVMEAPSQSL--PCVSEKPPQSHSDFLTQISRSPMLDCSIADCLQRCOV 1029
QY 1020 PFFGIQEEFNATLKNLFDWIKTSHNHLIVSTAEILFNDSTVTLPGQGAFFVRSQTE 1079
Db 1030 PSFSVQEEELDFTLKGLNLSFGWVRQILQKKVSVSVAEITFDTSVYSQLPQGEAFMRAQME 1089
QY 1080 TKVEPFEPVNPPLVTVGSSVGLLILALITAAALYKLGFFKQYKDMWSE 1128
Db 1090 MVLEDEEVNAPILMSSVGLLILALITATLYKLGFFKRYKEMLED 1138

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US20030077278A1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A COMB.pap: *
2: /cgn2_6/ptodata/1/iaa/5B COMB.pap: *
3: /cgn2_6/ptodata/1/iaa/6A COMB.pap: *
4: /cgn2_6/ptodata/1/iaa/6B COMB.pap: *
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap: *
6: /cgn2_6/ptodata/1/iaa/backfiles1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	1	US-08-173-497-3
2	5862	99.8	1153	1	US-08-286-889-3
3	5862	99.8	1153	1	US-08-485-618-3
4	5862	99.8	1153	1	US-08-362-652-3
5	5862	99.8	1153	2	US-08-605-672-3
6	5862	99.8	1153	2	US-08-482-293A-3
7	5862	99.8	1153	2	US-08-943-363-3
8	5862	99.8	1153	3	US-09-193-043-3
9	5862	99.8	1153	4	US-09-688-307A-3
10	5862	99.8	1153	4	US-09-350-259-3
11	5831.5	99.2	1152	2	US-08-476-062A-43
12	5831.5	99.2	1152	5	PCT-US96-01314-43
13	5831.5	99.2	1152	6	5424399-2
14	3469	59.0	1163	2	US-08-476-062A-44
15	3469	59.0	1163	5	PCT-US96-01314-44
16	3446	58.6	1163	1	US-08-173-497-4
17	3446	58.6	1163	1	US-08-286-889-4
18	3446	58.6	1163	1	US-08-485-618-4
19	3446	58.6	1163	1	US-08-362-652-4
20	3446	58.6	1163	2	US-08-605-672-4
21	3446	58.6	1163	2	US-08-482-293A-4
22	3446	58.6	1163	2	US-08-943-363-4
23	3446	58.6	1163	3	US-09-193-043-4
24	3446	58.6	1163	4	US-09-688-307A-4
25	3446	58.6	1163	4	US-09-350-259-4
26	3411	58.0	1161	1	US-08-173-497-2
27	3411	58.0	1161	1	US-08-286-889-2

28	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3411	58.0	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3411	58.0	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3411	58.0	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3411	58.0	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3411	58.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3411	58.0	1161	4	US-09-350-259-2	Sequence 99, Appl
36	3395.5	57.8	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3395.5	57.8	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3395.5	57.8	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3395.5	57.8	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3395.5	57.8	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3395.5	57.8	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3395.5	57.8	1161	4	US-09-350-259-99	Sequence 55, Appl
43	3232.5	55.0	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3232.5	55.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3232.5	55.0	1161	4	US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1

US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAFQEIIVAAHQGSLYQCDYSTGCEPI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAFQEIIVAAHQGSLYQCDYSTGCEPI 76

QY 61 RLQVPVAVNMVSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK 120
DB 77 RLQVPVAVNMVSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEOLKSKTFLSLMOYSEEF 196
QY 181 RIHFTTFEFOFNNPRSLIKPTQLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
DB 197 RIHFTTFEFOFNNPRSLIKPTQLGRTHATGIRKVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKESKFNNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPRYQHIGLVAMPR 420
DB 377 GVFLYTSKESKFNNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGTDLVLI GAPHYETRGQGVSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSDVDNSGTDLVLI GAPHYETRGQGVSVCP 496
QY 481 PRGORARWOCDAVLYGEQCPWGRFGAALTIVLDVNGDKLTVAICAPGEEDNRGAVYLF 540
DB 497 PRGORARWOCDAVLYGEQCPWGRFGAALTIVLDVNGDKLTVAICAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSORTAGSKLSPLOYFGOSLGGQDLTMDGLVDTVGAQGHVILLRSQ 600
DB 557 HGTSGSGISPSHSORTAGSKLSPLOYFGOSLGGQDLTMDGLVDTVGAQGHVILLRSQ 616
QY 601 PVLVRKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLREGIQSVVT 660
DB 617 PVLVRKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLREGIQSVVT 676
QY 661 YDLALDSGRPHSAVNETKSTRQTVLGLTCTETKLQLPNCIEDPVSIVLRLNF 720
DB 677 YDLALDSGRPHSAVNETKSTRQTVLGLTCTETKLQLPNCIEDPVSIVLRLNF 736
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMLDCLVVG 780
DB 737 SLVGTPLSAFNLRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVKSTLQNRQSRWRLACESASTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVKSTLQNRQSRWRLACESASTEV 856
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTSVSHGVSTKYNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVTSVSHGVSTKYNFTASENTSRVMQHYQVSNLQGRSLPISLVFLVPV 976
QY 961 RLNQTIVWDRPQVTFSENLSTCHTKERLPHSDFLAELRPAPVNCIAVCQRIQCDIP 1020
DB 977 RLNQTIVWDRPQVTFSENLSTCHTKERLPHSDFLAELRPAPVNCIAVCQRIQCDIP 1036
QY 1021 FPGIQEENATLKGNSFDWYIKTSHNLLIIVSTAEIILFNDVSFTLLPGQAGFVRSQTEF 1080
DB 1037 FPGIQEENATLKGNSFDWYIKTSHNLLIIVSTAEIILFNDVSFTLLPGQAGFVRSQTEF 1096
QY 1081 KVEPEVENPLPLIVSSVGGILLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1137
DB 1097 KVEPEVENPLPLIVSSVGGILLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1153

RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-286-889-3
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVAVNMVSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK 120
DB 77 RLQVPVAVNMVSLGLSLAATSPQALLACGPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEOLKSKTFLSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEOLKSKTFLSLMOYSEEF 196
QY 181 RIHFTTFEFOFNNPRSLIKPTQLGRTHATGIRKVVRELFTNGARKNAFKILILI 240
DB 197 RIHFTTFEFOFNNPRSLIKPTQLGRTHATGIRKVVRELFTNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPPRDHVFQIN 316
QY 301 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONLREKIFAIEGTQGTSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376


```
QY 361 GVFLYTSKEKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 420
Db 377 GVFLYTSKEKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496
QY 481 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRISAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRISAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 616
QY 601 PVLVRKVAIMEFNPVARNVFCNDQVYKKEAGEVRVCHLVQKSTRDLREGQIQSVVT 660
Db 617 PVLVRKVAIMEFNPVARNVFCNDQVYKKEAGEVRVCHLVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYVSNLQGRSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYVSNLQGRSLPISLVFLVPV 976
QY 961 RLNTQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020
Db 977 RLNTQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAGAFVRSQ 1080
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAGAFVRSQ 1096
QY 1081 KVEPFEVNPPLPIVSGVGLLALLITAAALYKLGFFKRYKQKDMMSGEGPPGAEPPQ 1137
Db 1097 KVEPFEVNPPLPIVSGVGLLALLITAAALYKLGFFKRYKQKDMMSGEGPPGAEPPQ 1153
```

RESULT 3

```
US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-485-618-3
```

```
Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQIIVAAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQIIVAAANORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFPLGSGNRQOPQK 120
Db 77 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGCLFPLGSGNRQOPQK 136
QY 121 FPALRGCPQEDSDIAFLIDVDSGSIIPHDFERRAKEFTVMEOLKSKTLFSLMOYSEEF 180
Db 137 FPALRGCPQEDSDIAFLIDVDSGSIIPHDFERRAKEFTVMEOLKSKTLFSLMOYSEEF 196
QY 181 RIHFTKFEQNNPNRSLIKPTITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RIHFTKFEQNNPNRSLIKPTITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPADREGVIRVVGCDAPRSEKSRQELNTVASKPPRDHVPQIN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRVVGCDAPRSEKSRQELNTVASKPPRDHVPQIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTSSESSFEHEMSQEGFSAATTSNGPILLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTSSESSFEHEMSQEGFSAATTSNGPILLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 436
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480
Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496
QY 481 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRQQRARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRISAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600
```

557 HGTSGSGISPSHSQRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTVGAGHVLRLRSQ 616
601 PVLRVKALMEFNPPEVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
617 PVLRVKALMEFNPPEVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQPNCTEDPSPVILRLNF 720
677 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQPNCTEDPSPVILRLNF 736
721 SLVGTPLSAFNLRLPVLAEADAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFNLRLPVLAEADAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900
857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916
901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTRVMOHQYQVSNLQORSPLISLVLPLVP 960
917 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTRVMOHQYQVSNLQORSPLISLVLPLVP 976
961 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPDSHDFLAELRKAPVNVNCSIAVCORIQCDIP 1020
977 RLNQTVIWDPRQVTFSENLSSCTCHTKERLPDSHDFLAELRKAPVNVNCSIAVCORIQCDIP 1036
1021 PFGIOEEFNATLKNLSFOWYIKTSNHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080
1037 PFGIOEEFNATLKNLSFOWYIKTSNHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1096
1081 KVEPFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFRKQYKDMSEGGPGCAEPQ 1137
1097 KVEPFEPVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFRKQYKDMSEGGPGCAEPQ 1153

RESULT 4
US-08-362-652-3
; Sequence 3, Application US/08362652
; Patent No. 5766850
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,652
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGTVHOTCSENTVVKGLCFLFGSNLRQOPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGTVHOTCSENTVVKGLCFLFGSNLRQOPOK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKKSKTLFSLMOYSBEP 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKKSKTLFSLMOYSBEP 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTNGARKNAFKLLIL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTNGARKNAFKLLIL 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAPFSEKSRQELNNTVASKPRDHVQFN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAPFSEKSRQELNNTVASKPRDHVQFN 316
QY 301 NFEALKTIONOLREKIFAIEGTQTSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKIFAIEGTQTSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIIILNRVQSLVLGAPRYOHLGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIIILNRVQSLVLGAPRYOHLGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYVEQTRGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYVEQTRGGQVSCPL 496
QY 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 540
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 556
QY 541 HGTSGSGISPSHSQRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTVGAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTVGAGHVLRLRSQ 616
QY 601 PVLRVKALMEFNPPEVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660
Db 617 PVLRVKALMEFNPPEVARNVPECDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQPNCTEDPSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQPNCTEDPSPVILRLNF 736
QY 721 SLVGTPLSAFNLRLPVLAEADAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFNLRLPVLAEADAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENNMPTNKTEF 900
Db 857 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYQVSNLQSRSLPSLFLVFPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYQVSNLQSRSLPSLFLVFPV 976
QY 961 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036
QY 1021 PFGIOEFNATLKGNSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080
Db 1037 PFGIOEFNATLKGNSPDWYIKTSHNHLIIYSTAEILFNDVSFTLLPGOGAFVRSQTEF 1096
QY 1081 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEFPQ 1137
Db 1097 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMWSEGGPPGAEFPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817315el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENLDTENAMTFOENARGFCQSVVQLQSGSRVVGAPQEIIVAANQRGSLYQCDYSGSCEPI 60
Db 17 FNLDTENAMTFOENARGFCQSVVQLQSGSRVVGAPQEIIVAANQRGSLYQCDYSGSCEPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFIPTVMEQLKKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFIPTVMEQLKKSKTLFSLMQYSEEF 196
QY 181 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGIRKVVRELFTNITNGARKNAFKILVI 240
Db 197 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGIRKVVRELFTNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVFQIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPRDRHVFQIN 316
QY 301 NFEALKTTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
QY 361 GVELYTSKEKSTFINMTFVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTFVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTEGGQVSCPL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTEGGQVSCPL 496
QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNNGAVYLF 540
Db 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNNGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTVGAGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTVGAGHVLLLRSQ 616
QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTDRLRREGIOISVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTDRLRREGIOISVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNTSTRRTQVLGLTQTCETLKLQFNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSRAVFNETKNTSTRRTQVLGLTQTCETLKLQFNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPEKNCNDNICQDDLSITFSFMSLDCLVWG 796
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENNMPTNKTEF 900
Db 857 SGALKSTCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYQVSNLQSRSLPSLFLVFPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQYQVSNLQSRSLPSLFLVFPV 976
QY 961 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036

1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
1081 KVEPFEVPNPLPLVGVSGVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
1097 KVEPFEVPNPLPLVGVSGVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 6
US-08-482-293A-3
; Sequence 3, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994

ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

1 FNLDTENAMTPOENARGFGQSVVQLQGSRRVVVGAPOEIVANQSGSLYQCDYSTGCEPI 60
17 FNLDTENAMTPOENARGFGQSVVQLQGSRRVVVGAPOEIVANQSGSLYQCDYSTGCEPI 76

61 RLQPVAVNMSLGLSLAATTSPPOLLACGFTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
77 RLQPVAVNMSLGLSLAATTSPPOLLACGFTVHQTCSNTYVKGICFLFGSNLRQOPQK 136

121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAXEFTSTVMEQLKSKTLFSLMQYSEEF 180

137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAXEFTSTVMEQLKSKTLFSLMQYSEEF 196
181 RIHFTFRFQNNPRSLRPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
197 RIHFTFRFQNNPRSLRPIITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVPEADREGVIRVIVGVDAPFRSEKSEQLNTVASKPRPHVPOIN 300
257 TDGEKFGDPLGYEDVPEADREGVIRVIVGVDAPFRSEKSEQLNTIASKEPRPHVPOIN 316
301 NFEALKTIQNLREKIIFAEIGTQSGSSSFHEMSQEGFSAITNSGILLTSTVSYDWAG 360
317 NFEALKTIQNLREKIIFAEIGTQSGSSSFHEMSQEGFSAITNSGILLTSTVSYDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGRGQSVVCP 480
437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTGRGQSVVCP 496
481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 540
497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 556
541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVLTVAQGHVLLLRSQ 600
557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVLTVAQGHVLLLRSQ 616
601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVVCLHVQKSTRDLRGQIQSVVT 660
617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVVCLHVQKSTRDLRGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKSTRQTVLGTQTCETLKLQPNCEIDPVPVILRLNF 720
677 YDLALDSGRPHSRVAFNETKSTRQTVLGTQTCETLKLQPNCEIDPVPVILRLNF 736
721 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFKEKNGNDNICODDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFKEKNGNDNICODDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRQSWLACESASSTEV 840
797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRQSWLACESASSTEV 856
841 SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
857 SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 916
901 QLELPVKYAVYVMTVSHGVSTKYINFNTASNTSRVMQHVQVSNLQGRSLPISLVFLVPV 960
917 QLELPVKYAVYVMTVSHGVSTKYINFNTASNTSRVMQHVQVSNLQGRSLPISLVFLVPV 976
961 RLNQTVIWRDPQVTFSENTSSTCHTKERLPDLSHSDFLAELKAPVWNCIAVCORICDIP 1020
977 RLNQTVIWRDPQVTFSENTSSTCHTKERLPDLSHSDFLAELKAPVWNCIAVCORICDIP 1036
1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
1081 KVEPFEVPNPLPLVGVSGVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
1097 KVEPFEVPNPLPLVGVSGVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 7
US-08-943-363-3
; Sequence 3, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
 APPLICANT: Van der Vieren, Monica
 TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
 NUMBER OF SEQUENCES: 114
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshhall, O'Toole, Gerstein, Murray & Borun
 STREET: 233 South Wacker Drive, 6300 Sear Tower
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/943,363
 FILING DATE: 5-AUG-1994

CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/173,497
 FILING DATE: 23-DEC-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/286,889
 FILING DATE: 5-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/362,652
 FILING DATE: 21-DEC-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Williams Jr., Joseph A.
 REGISTRATION NUMBER: 38,659
 REFERENCE/DOCKET NUMBER: 27866/32684
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312-474-6300
 TELEFAX: 312-474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1153 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-943-363-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVANORGSLYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVANORGSLYQCDYSTGSCPEI	76
QY	61	RLOQVPEAVNMSLGLSLAATTSPPQLIACGPTVHQTCSENTYVKGCLFLGNSLRQOPQK	120
Db	77	RLOQVPEAVNMSLGLSLAATTSPPQLIACGPTVHQTCSENTYVKGCLFLGNSLRQOPQK	136
QY	121	FPEALRGCPQEDSDIAFLVDGSGIIPHPDRAKEPISTVMEQLKSKTLFSLMQYSEEF	180
Db	137	FPEALRGCPQEDSDIAFLVDGSGIIPHPDRAKEPISTVMEQLKSKTLFSLMQYSEEF	196
QY	181	RIHFTKFEQNNPNSRLKIPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIT	240
Db	197	RIHFTKFEQNNPNSRLKIPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIT	256
QY	241	TGKEKGDPLGVEDVTPEDADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVQFN	300
Db	257	TGKEKGDPLGVEDVTPEDADREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVQFN	316
QY	301	NFEALKTIONQRLREKIFATEGTQTGSSSFEHMSQEGFSAITSNGPLLSTVGSYDMAG	360
Db	317	NFEALKTIONQRLREKIFATEGTQTGSSSFEHMSQEGFSAITSNGPLLSTVGSYDMAG	376

QY	361	GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMFR	420
Db	377	GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMFR	436
QY	421	QNTGMWESNANVKGTOIGAYFGASICSVDVDNSGSTDVLIIGAPHYYEOTRGQSVCP	480
Db	437	QNTGMWESNANVKGTOIGAYFGASICSVDVDNSGSTDVLIIGAPHYYEOTRGQSVCP	496
QY	481	PRGQARWQCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAITGAPGEEDNRGAVLYF	540
Db	497	PRGQARWQCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDAITGAPGEEDNRGAVLYF	556
QY	541	HGTSGSIGSPSHSQRISQISKLSPRLQYFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ	600
Db	557	HGTSGSIGSPSHSQRISQISKLSPRLQYFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ	616
QY	601	PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	617	PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSRAVFNETKNSRRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF	720
Db	677	YDLALDSGRPHSRAVFNETKNSRRQTOVLGLTQTCETLKLQLPNCIEDPVPVLRNLF	736
QY	721	SLVGTPLSAFONLRPVLAEADQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG	780
Db	737	SLVGTPLSAFONLRPVLAEADQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG	796
QY	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV	840
Db	797	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV	856
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF	900
Db	857	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF	916
QY	901	QLELPVKYAVYVMVTSHGVTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	960
Db	917	QLELPVKYAVYVMVTSHGVTKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	976
QY	961	RLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQRIQDIP	1020
Db	977	RLNQTIVWDRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQRIQDIP	1036
QY	1021	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETLNDVSFTLLPGOGAFVRSQTET	1080
Db	1037	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETLNDVSFTLLPGOGAFVRSQTET	1096
QY	1081	KVEPFVFPNPLPIVGVSSVGLLALLALITAALYKLGFFKQYKDMKSEGGPPGAEPQ	1137
Db	1097	KVEPFVFPNPLPIVGVSSVGLLALLALITAALYKLGFFKQYKDMKSEGGPPGAEPQ	1153

RESULT 8

US-09-193-043-3
 ; Sequence 3, Application US/09193043
 ; Patent No. 6251395
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, Michael W.
 ; APPLICANT: Van der Vieren, Monica
 ; TITLE OF INVENTION: No. 6251395e1 Human 2
 ; FILE REFERENCE: 27866/35004
 ; CURRENT APPLICATION NUMBER: US/09/193,043
 ; CURRENT FILING DATE: 1998-11-16
 ; EARLIER APPLICATION NUMBER: 08/173,497
 ; EARLIER FILING DATE: 1993-12-23
 ; EARLIER APPLICATION NUMBER: 08/286,889
 ; EARLIER FILING DATE: 1994-08-05
 ; EARLIER APPLICATION NUMBER: 08/362,652
 ; EARLIER FILING DATE: 1994-12-21
 ; EARLIER APPLICATION NUMBER: 08/943,363
 ; EARLIER FILING DATE: 1997-10-03

Db 317 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAITSNGLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVCP 480
Db 437 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVCP 496
Qy 481 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
Qy 661 YDLALDSGRPHSAVNETKNSTRQTQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSAVNETKNSTRQTQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736
Qy 721 SLVGTPLSAFNLPRVLAEDAQRLLFTALPFEKNCNDNIQDDLSITFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNLPRVLAEDAQRLLFTALPFEKNCNDNIQDDLSITFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWLACESASSTEV 856
Qy 841 SGALKSTCSINHPFPENSEVTFNTFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTCSINHPFPENSEVTFNTFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
Qy 901 QLELPVKAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 960
Db 917 QLELPVKAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV 976
Qy 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPHSDPLAELRKAAPVNCVSIQVQRIQCDIP 1020
Db 977 RLNQTVIWDROPVTFSENLSSTCHTKERLPHSDPLAELRKAAPVNCVSIQVQRIQCDIP 1036
Qy 1021 PFGIQEEFNATLKGNSLFDWIKTSHNHLIVSTAELFNDVSVFTLLPQOGAFVRSQTET 1080
Db 1037 PFGIQEEFNATLKGNSLFDWIKTSHNHLIVSTAELFNDVSVFTLLPQOGAFVRSQTET 1096
Qy 1081 KVEPFEVNPPLIVGSSVGLLALLALITAALYKLGFFKRYKQKDMSEGGPPGABPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGLLALLALITAALYKLGFFKRYKQKDMSEGGPPGABPQ 1153

RESULT 10

US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVCAPOEIVAAORGSLYQCDYGTGCEPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVCAPOEIVAAORGSLYQCDYGTGCEPI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGLCFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQTCSENTYVKGLCFLFGSNLRQPOK 136
Qy 121 EPEALRGCPQEDSDIAFLVDGSGSIIPHPDRAKEPISTWMEQLKKSKTLFSLMOYSEBF 180
Db 137 EPEALRGCPQEDSDIAFLVDGSGSIIPHPDRAKEPISTWMEQLKKSKTLFSLMOYSEBF 196
Qy 181 RHFTFKFQONNPNRSLIKPITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RHFTFKFQONNPNRSLIKPITQLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
Qy 241 TDGEXKPGDPLGYEDVIPEADREGVIRYVIGVGDADFSEKSRQBLNVTASKPRDHVFOIN 300
Db 257 TDGEXKPGDPLGYEDVIPEADREGVIRYVIGVGDADFSEKSRQBLNVTASKPRDHVFOIN 316
Qy 301 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAITSNGLSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAITSNGLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVCP 480
Db 437 QNTGWSNANVKTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQSVCP 496
Qy 481 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSORIASGKSLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGKSLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 616
Qy 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
Qy 661 YDLALDSGRPHSAVNETKNSTRQTQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSAVNETKNSTRQTQVGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736
Qy 721 SLVGTPLSAFNLPRVLAEDAQRLLFTALPFEKNCNDNIQDDLSITFSMSLDCLVVG 780
Db 737 SLVGTPLSAFNLPRVLAEDAQRLLFTALPFEKNCNDNIQDDLSITFSMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWLACESASSTEV 856
Qy 841 SGALKSTCSINHPFPENSEVTFNTFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900

857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQKSLPISLVFLVPV 960
Db 917 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQKSLPISLVFLVPV 976
QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCORIQCDIP 1036
QY 1021 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1080
Db 1037 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1096
QY 1081 KVEPEVENPLPLIVGSSVGGLLLLALITALYKLGFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPEVENPLPLIVGSSVGGLLLLALITALYKLGFFKQYKDMMSGGPPGABPQ 1153

RESULT 11
US-08-476-062A-43
; Sequence 43, Application US/08476062A
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELETYPE: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: Internal
; US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

1 FNLDTENAMTFQENARFGQSVVOLQGRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVOLQGRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENTYVKGCLFCLFSGNLRRQDPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENTYVKGCLFCLFSGNLRRQDPQK 136
QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRAKEFIISTVMEOCLKSKTSLSLMOYSSEF 180
Db 137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRMKEFVSTVMEQCLKSKTSLSLMOYSSEF 196
QY 181 RIHFTTFKEFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
Db 197 RIHFTTFKEFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELMTVASKPRDHVFOIN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELMTVASKPRDHVFOIN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLLSTVSGYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLLSTVSGYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAIIILNRVQSLVIGAPYQHQHGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAIIILNRVQSLVIGAPYQHQHGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFCASLCSVDVDSNGSTDLVLIGAPHYVEQTRGQSVVCEL 480
Db 437 QNTGMWESNANVKGTQIGAYFCASLCSVDVDSNGSTDLVLIGAPHYVEQTRGQSVVCEL 496
QY 481 PRGORARWOCDAVLYGEOQGPWGRGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAYILF 540
Db 497 PRG-RARWOCDAVLYGEOQGPWGRGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAYILF 555
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLLRSQ 615
QY 601 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 616 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRNF 720
Db 676 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRNF 735
QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALPFEKNCNDNICODDLSITPFSMLDCLVVG 780
Db 736 SLVGTPLSAFNLRPVLAEDAQRLLFTALPFEKNCNDNICODDLSITPFSMLDCLVVG 795
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQKSLPISLVFLVPV 960
Db 916 QLELPVKYAVYVMVTSYKYNFTASNTSRVMOHQYQVSNLQKSLPISLVFLVPV 975
QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCORIQCDIP 1020
Db 976 RLNOTVIWDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCORIQCDIP 1035
QY 1021 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1080
Db 1036 FFGIOEEFNATLKGNLSPDWYIKTSHNHLIVSTAEIILFNDSVFTLLPQCGAFVRSQDET 1095
QY 1081 KVEPEVENPLPLIVGSSVGGLLLLALITALYKLGFFKQYKDMMSGGPPGABPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 12

PCT-US96-01314-43

Sequence 43, Application PC/TUS9601314

GENERAL INFORMATION:

APPLICANT: M. Amin Arnaut

TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN

NUMBER OF SEQUENCES: 78

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 50Z or 55SX

OPERATING SYSTEM: MS-DOS (Version 5.0)

SOFTWARE: WordPerfect (Version 5.1)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/01314

FILING DATE: 30-JAN-96

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/380,167

FILING DATE: 30-JAN-95

ATTORNEY/AGENT INFORMATION:

NAME: John W. Freeman

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 00786/267001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 1152

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

PCT-US96-01314-43

Query Match 99.2%; Score 5831.5; DB 5; Length 1152;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQIIVANQSGSLYQCDYSTGSCPEI 60

Db 17 FNLDTENAMTFOENARGFGOSVVOLOGSRVVGAPQIIVANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLROQPOK 120

Db 77 RLQVPEAVNMSGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLROQPOK 136

QY 121 FPALRGCPQEDSDIAPLVGSGSIIPHDFRAKEPISTVMEQIKKSKTLFSLMQYSEEF 180

Db 137 FPALRGCPQEDSDIAPLVGSGSIIPHDFRMRKEFVSTVMEQIKKSKTLFSLMQYSEEF 196

QY 181 RIHFTKBFQNNPRSLKPIITQLGRTHATGIRKVVRELNITNGARKNAFKILVLI 240

Db 197 RIHFTKBFQNNPRSLKPIITQLGRTHATGIRKVVRELNITNGARKNAFKILVLI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 300

Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRQELNTVASKPPRDHVFQIN 316

QY 301 NFEALKTIONQREKIFAIEGTGSSSFEHEMSQEGFSAITSNGLPSTVGSYDWMAG 360

Db 317 NFEALKTIONQREKIFAIEGTGSSSFEHEMSQEGFSAITSNGLPSTVGSYDWMAG 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILLNRVQSLVGLAPRHOHIGLVAMFR 420

Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAALILLNRVQSLVGLAPRHOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 480

Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 496

QY 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYL 540

Db 497 PRG-RARWOCDAVLYGEOGQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYL 555

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 600

Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 615

QY 601 PVLRYKAIMFNPREVARNVFECDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 660

Db 616 PVLRYKAIMFNPREVARNVFECDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 675

QY 661 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQIPNCIEDPVPPIVRLNMF 720

Db 676 YDLALDSGRPHSRVAVNETKSTRQTOVLGLTQTCETLKLQIPNCIEDPVPPIVRLNMF 735

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFPEKNGCNDNICODLSITFSEMSLCLVVG 780

Db 736 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFPEKNGCNDNICODLSITFSEMSLCLVVG 795

QY 781 GPRENVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 840

Db 796 GPRENVTVTVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWELACESASSTEV 855

QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900

Db 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 915

QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHVOYVSNLQSRSLPISLVFLVPV 960

Db 916 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHVOYVSNLQSRSLPISLVFLVPV 975

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020

Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035

QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETILFNDVSFTLLPGQAFVRSQTE 1080

Db 1036 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETILFNDVSFTLLPGQAFVRSQTE 1095

QY 1081 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1152

RESULT 13

5424399-2

Patent No. 5424399

APPLICANT: ARNAOUT, M. AMIN

TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS

NUMBER OF SEQUENCES: 12

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/78,871

FILING DATE: 16-JUN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 539,842

FILING DATE: 18-JUN-1990

APPLICATION NUMBER: 212,573

FILING DATE: 28-JUN-1988

SEQ ID NO: 2;

LENGTH: 1152

5424399-2

Query Match 99.2%; Score 5831.5; DB 6; Length 1152;

1036 FFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQGAFFVRSQTE 1095

1081 KVPFFVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1137
1096 KVPFFVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1152

RESULT 14

US-08-476-062A-44
; Sequence 44: Application US/08476062A
; Patent No. 5877275
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: Fast-SEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,062A
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/216,081
; FILING DATE: 21-MAR-1994
; APPLICATION NUMBER: 07/637,830
; FILING DATE: 04-JAN-1991
; APPLICATION NUMBER: 07/539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 07/212,573
; FILING DATE: 28-JUN-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/068003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1163 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-476-062A-44

Query Match 59.0%; Score 3469; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 7e-284;
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVVGAPOEIVANORGLYQCDYSTGSCPEI 60
Db 20 FNLDTELTAFRVDGSGSVVQYANVSVVVGAPQKLTAAQNGGLYQCGYSTGACEPI 79
Qy 61 RLQVPVAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQBPQK 120
Db 80 GLQVPPVAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQBPQK 137
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPPRAKEFTSTVMEQLKSKTLFSLMOYSSEF 180
Db 138 LPVSRQECPCREQDQIVFLIDGSGSISRNFNATMMNFVRAVISQFORPSTQFSLMOFSNKF 197
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFOENARFGQSVVQLQGSRRVVVGAPOEIVANORGLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARFGQSVVQLQGSRRVVVGAPOEIVANORGLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQBPQK 120
Db 77 RLQVPPVAVNMVSLGLSLAATTPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQBPQK 136
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPPRAKEFTSTVMEQLKSKTLFSLMOYSSEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPPRAKEFTSTVMEQLKSKTLFSLMOYSSEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILIVI 256
Qy 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKROELNTVASKPPRDHVFQIN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKROELNTIASKPPRDHVFQIN 316
Qy 301 NFEALKTQONLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAQ 360
Db 317 NFEALKTQONLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAQ 376
Qy 361 GYFLVTSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPYQHIGLVAMPR 420
Db 377 GYFLVTSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480
Db 437 QNTGMWESNANVKGTOIGAFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496
Qy 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLDVNGDKLTDVAICAPGEEDNRGAVYLF 540
Db 497 PRG-RARWQCDVLYGEGQGWGRFGAALTIVLDVNGDKLTDVAICAPGEEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSORLAGSKSLRLOVFGQSLGGQDLTMDGLVDLTVGAGHVLLRSQ 600
Db 556 HGTSGSGISPSHSORLAGSKSLRLOVFGQSLGGQDLTMDGLVDLTVGAGHVLLRSQ 615
Qy 601 PVLVRKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 616 PVLVRKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 675
Qy 661 YDLALDGRPHSRVAFNETKNSTRQTOVLGLTQCTETLKLQPNCLIEDPVSPVLRNLF 720
Db 676 YDLALDGRPHSRVAFNETKNSTRQTOVLGLTQCTETLKLQPNCLIEDPVSPVLRNLF 735
Qy 721 SLVCTPLSAFGLNRPVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 SLVCTPLSAFGLNRPVLAEDAQRLLFTALFFPEKNCNDNICQDDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGDSRYRTQVTFPPPLDLSYKXVSTLQVRSORSWRLACESASSTEV 840
Db 796 GPREFNVTVVRNDGDSRYRTQVTFPPPLDLSYKXVSTLQVRSORSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNPNRPNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNPNRPNKTEF 915
Qy 901 QLELPVKYAVVMVTSVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSPIVFLVPV 960
Db 916 QLELPVKYAVVMVTSVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQORSLSPIVFLVPV 975
Qy 961 RLNQTIVWDRPOVTFSENLSTCTHTKRLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
Db 976 RLNQTIVWDRPOVTFSENLSTCTHTKRLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1035
Qy 1021 FFGIQEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQGAFFVRSQTE 1080

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds
(without alignments)
4014.622 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	1 RWHU1B	cell surface glyco
2	4476	76.2	1153	2 S00551	leukocyte surface
3	3483	59.3	1163	1 RWHU1C	cell surface glyco
4	1548.5	26.4	1170	2 S03308	cell surface glyco
5	1533.5	26.1	1163	2 I56126	lymphocyte fuction
6	1149	19.6	1179	2 A53213	integrin alpha-E c
7	1102.5	18.8	1151	2 A43226	integrin alpha-1 c
8	1084	18.5	1170	2 I45914	integrin alpha-1 c
9	1072	18.2	1178	2 S44142	VLA-2 protein homo
10	1069	18.2	1181	2 A33998	integrin alpha-2 c
11	1060	18.0	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	633	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-4 c
15	614.5	10.5	1041	2 T34437	integrin alpha-9 c
16	579.5	9.9	1054	2 JC7294	integrin alpha-cha
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	543.5	9.3	1034	2 A36108	integrin alpha-V c
21	535	9.1	1049	2 T10050	integrin alpha-V c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	530.5	9.0	1072	2 A38457	integrin alpha-6 c
25	529.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.5	1146	2 S40311	integrin - fruit f

RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain
C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A36091; I52567

R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b).

A;Reference number: A31108; MUID:88315033; PMID:2457584

A;Accession: A31108

A;Molecule type: mRNA

A;Residues: 1-1153 <COR>

A;Cross-references: UNIPROT:P11215; GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:g307148

A;Note: part of this sequence was confirmed by protein sequencing

R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor N

A;Reference number: A28915; MUID:88257215; PMID:2454931

A;Accession: A28915

A;Molecule type: mRNA

A;Residues: 1-499,501-965,'P',967-1153 <ARN>

A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA594

A;Note: the authors translated the codon TAC for residue 1129 as Thr

A;Note: part of this sequence, including the amino end of the mature protein, was confir

R;Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A;Reference number: A41600; MUID:92073318; PMID:1683702

A;Accession: A41600

A;Molecule type: DNA

A;Residues: 1-9 <SHB>

A;Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:g553215

R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes

A;Accession: A30892

A;Molecule type: mRNA

A;Residues: 917-1042 <AR2>

A;Cross-references: GB:M18044

R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A;Title: CDNA sequence for the alphaM subunit of the human neutrophil adherence recepto

A;Reference number: A32218; MUID:89098893; PMID:2563162

A;Accession: A32218

A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>

A;Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:g386975

integrin alpha-2b
position-specific
integrin alpha-7 c
alpha-7 integrin -
leukocyte adhesion
glycoprotein IIB -
integrin alpha cha
F54F2.1 protein -
integrin alpha v c
hypothetical prote
integrin alpha cha
integrin alpha cha
glycoprotein IIB
integrin alpha cha
glycoprotein IIB -
integrin alpha-1 -

ALIGNMENTS

A;Note: part of this sequence was confirmed by protein sequencing
 R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.
 J. Immunol. 150, 480-490, 1993
 A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in
 n during evolution.
 A;Reference number: A46526; MUID:93123748; PMID:8419480
 A;Accession: A46526
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-499,501-1153 <FILE>
 A;Cross-references: GB:552227; NID:g263047; PIDN:AA24821.1; PID:g263049
 A;Note: the last three bases of intron 13, CAG, are included in some but not all mature
 A;Note: sequence extracted from NCBI backbone (NCBIP:121963)
 R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.
 Biochim. Biophys. Acta 874, 368-371, 1986
 A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp
 A;Reference number: A90664; MUID:87076671; PMID:3539202
 A;Accession: A26091
 A;Molecule type: protein
 A;Residues: 17-31 <PIE>
 A;Experimental source: granulocytes
 R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.
 Blood 79, 865-870, 1992
 A;Title: Characterization of the myeloid-specific CD11b promoter.
 A;Reference number: 152567; MUID:12144986; PMID:1346576
 A;Accession: 152567
 A;Status: translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-9 <RES>
 A;Cross-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:g553219
 C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1
 C;Genetics:
 A;Gene: GDB:ITGAM; CR3A
 A;Cross-references: GDB:120599; OMIM:120980
 A;Map position: 16p11.2-16p11.2
 A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
 C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag
 F;1-16/Domain: signal sequence #status predicted <SIG>
 F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>
 F;17-1108/Domain: extracellular #status predicted <EXT>
 F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>
 F;465-473/Region: calcium/magnesium binding #status predicted
 F;530-538/Region: calcium/magnesium binding #status predicted
 F;593-601/Region: calcium/magnesium binding #status predicted
 F;1109-1113/Domain: transmembrane #status predicted <TM>
 F;1135-1153/Domain: intracellular #status predicted <INT>
 F;86,240,391,469,693,797,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

QY 301 NFALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFAAITSNGPLLSITVGSYDWAG 360
 Db |||||
 QY 317 NFALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFAAITSNGPLLSITVGSYDWAG 376
 Db |||||
 QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 420
 Db |||||
 QY 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHIGLVAMFR 436
 Db |||||
 QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEYQTRGGQSVCP 480
 Db |||||
 QY 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEYQTRGGQSVCP 496
 Db |||||
 QY 481 PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540
 Db |||||
 QY 497 PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 556
 Db |||||
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLLR 600
 Db |||||
 QY 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLLR 616
 Db |||||
 QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
 Db |||||
 QY 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
 Db |||||
 QY 661 YDLALDSGRPHSRAVFNETKNSTRROTQVLGLTQTCTETKLQLPNCIEDPVSPIVLRNF 720
 Db |||||
 QY 677 YDLALDSGRPHSRAVFNETKNSTRROTQVLGLTQTCTETKLQLPNCIEDPVSPIVLRNF 736
 Db |||||
 QY 721 SLVGTPLSAFGNLRPVLAEADQRLFTALFPFKNCGNDNICDDLSITFSFMSLDCLVVG 780
 Db |||||
 QY 737 SLVGTPLSAFGNLRPVLAEADQRLFTALFPFKNCGNDNICDDLSITFSFMSLDCLVVG 796
 Db |||||
 QY 781 GPREFNVTVTRNDGDSYETQVTFPEPLDLSVRKYSTLONORSORSMWLACESASSTEV 840
 Db |||||
 QY 797 GPREFNVTVTRNDGDSYETQVTFPEPLDLSVRKYSTLONORSORSMWLACESASSTEV 856
 Db |||||
 QY 841 SGALKSTSCSINHPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 Db |||||
 QY 857 SGALKSTSCSINHPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
 Db |||||
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHQVSNLQORSPIPLSVFLVPV 960
 Db |||||
 QY 917 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHQVSNLQORSPIPLSVFLVPV 976
 Db |||||
 QY 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSSHDFLAELKAPVNVCSIAVCQRIQCDIP 1020
 Db |||||
 QY 977 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSSHDFLAELKAPVNVCSIAVCQRIQCDIP 1036
 Db |||||
 QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080
 Db |||||
 QY 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096
 Db |||||
 QY 1081 KVEPFEPVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKQDMSEGPPGAEQ 1137
 Db |||||
 QY 1097 KVEPFEPVNPPLIVGSSVGLLALLITAAALYKLGFFKQYKQDMSEGPPGAEQ 1153
 Db |||||

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C;Accession: S00551; 159078

R;Pyteia, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

A;Molecule type: DNA

A;Residues: 1-1153 <PYT>

A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:952982; PIDN:CAA30479.1; PID:95298

A;Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep
A:Reference number: 159078; MUID:86287312; PMID:12942940
A:Accession: 159078
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 11-44 <RES>
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193
C:Genetics:
A:Gene: Mac-1
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: cell adhesion; glycoprotein; transmembrane protein
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental
F:148-318/Domain: von Willebrand factor type A repeat homology <WMA>
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.2%; Score 4476; DB 2; Length 1153;
Best Local Similarity 74.3%; Pred. No. 6.1e-300;
Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFOENARFGSGVOLQSGRVVVGAPQEIIVANQSGISYQCDYSTGSCPEI 60
Db 17 FNLDTEHPMTFOENAKGFGQNVQLGTSVVAAPQBAKAVNOTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSGLSLAATTSPPOLACGPTVHOTCSENTYVYKGLCFLEGSNLRQPOK 120
Db 77 PLQVPEAVNMSGLSLAVSTVPQQLLACGPTVHONCKENTYVNGLCYLFSGNLLRPPQ 136

QY 121 FPEALRGCPQEDSDIAPLIDGSGSIIIPDFRRMKEFVSTVMEOLKSKTILFSLMQSSEF 180
Db 137 FPEALRGCPQESDIPLIDGSGSINNIDFQMKKEFVSTVMEQFKSKTILFSLMQSDEF 196

QY 181 RIHTFKFQNNNPSRLVKPIITQLGRHTATGVRKVRIBELNITNGARKNAFKILIVI 240
Db 197 RIHTFNDFRNPSRSHVSPKQLNGRTKTASGIRKVRVRELFHKTNGARENAKILIVI 256

QY 241 TDEKFGDPLGYEDVPEADREGVIRVIGVGFAPRSEKROBLNTIASPPRDHVPQVN 300
Db 257 TDEKFGDPLGYEDVPEADREGVIRVIGVGNAPRSEKRLDITIASPAGEHFQVD 316

QY 301 NFBAKTIQNLREKIFAIBGTQTGSSSPSEHMSQEGFSAALTNGPLSLTVSGVDWAG 360
Db 317 NFBAKTIQNLREKIFAIBGTQTGSSSPSEHMSQEGFSAALTNGPLSLTVSGVDWAG 376

QY 361 GVFLYTSKEKSTPINTRVDSMDNDAYLVGAAILRNVRQSVILGAPRYOHIGLVAMFR 420
Db 377 GAFLYTSKOKVTPINTTRVDSMDNDAYLVGASAVILRNVRQSVILGAPRYOHIGLVAMFR 436

QY 421 QNTGMESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCP 480
Db 437 ENFGTWEPTSIAGSQIGSTYFGASLCSVDMDAGNTNLIIGAPHYYEOTRGQSVCP 496

QY 481 PRQARWQCDAYLYEQGQFQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNKGAVYLF 540
Db 497 PRG-RARWQCEALLHGDQGHQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEQNGQAVYIF 555

QY 541 HGTSGSGISPSHSQRTAGSKLSRPLYFGQSLSGGQDLTMDGLVDITVGAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIGAHFSPGLQVFGQSLSGGKDLTMDGLMDLAVGAQGHVLLRAQ 615

QY 601 PVLVRVAIMEFNPREVARNVECDQVKGKEAGEVRVCLHVOKSTRDLRREGIQSVVT 660
Db 616 PVLRLSATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLVRKNTKDLRREGDIQSTVT 675

QY 661 YDLALDSGRPHSAFVNETKNSRRTOTVGLTQTCTETLKLQPCNTEDPVSPVILRLNF 720
Db 676 YDLALDPVSRIRAPFDETKNTRRTQVFLGMQKQKTELKLIIDPCVDVSPVILRLNY 735

QY 721 SLVGTPLSAFGLRPVLAEDAQRALTALFPFPEKNCNDNICDDLSITFSMSLCLVVG 780
Db 736 TLVGEPLRSGNLRPVLAMDQARFFTAFFPEKNCNDNICDDLSITFSMSGLDLTVVG 795

QY 781 GPREFNVTVRNDGEDSYQTQVTFPPPLDLSVYKVKSTLQNRQSRWRLL-ACSAASSTE 839
Db 796 GPQDFNMSVTLRNDGEDSYQTQVTVYVPSGLSVKRDASQNPITTKKPFVFKPAESSSSSE 855

QY 840 VSGALKSTSGSINHPIPPENSEVENTITFDVDSKASIGNKLLKANVTSENMMPTNKTE 899
Db 856 GHGALKSTTWNINHPFPANSEVTFNFDVDSHSPGNKLLKAI VASENNMERTHAKTK 915

QY 900 FQLELPVKYAVYVMTSHGVSTKYLNFTASENSTRVQMHOQVQVSNLQGRSLPISLVFLVP 959
Db 916 FQLELPVKYALYIMVTSDESSIRYLNFTASEMTSKVQHQQYNLQGRSLPVSVVFWIP 975

QY 960 VRLNQTVINDRPOQVTSSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDI 1019
Db 976 VQINVTVDHPQVIFESQNLSSACHTEQKSPPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035

QY 1020 PFTGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAILFNDSVFTLLPGGAFVRSOTE 1079
Db 1036 PSFTQEIFNVTILKNLSFDWYIKTSHGHLULLSVSTELFNDSAFALLPQGESVRSKTE 1095

QY 1080 TKVEPEVNPPLPLIVGSSVGGLLALLALITAAALYKLGFFKRYQKDMMSGPPGABPQ 1137
Db 1096 TKVEPEVHNPPVLLIVGSSIGGLVALLALITAGLYKLGFFKRYQKDMWNEAAPQDAPQ 1153

RESULT 3
RWHLUC
cell surface glycoprotein CD11c precursor - human
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36584; A35543; S00864
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 12750-12751, 1990
A:Reference number: A36584
A:Contents: erratum
A:Accession: A36584
A:Molecule type: DNA
A:Residues: 1-1163 <COR>
A:Cross-references: UNIPROT:P20702
A:Note: this revision to the sequence from reference A35543 includes the carboxyl end
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule
A:Reference number: A35543; MUID:90153906; PMID:2303426
A:Accession: A35543
A:Molecule type: DNA
A:Residues: 1-834 <CO2>
A:Note: this sequence has been revised in reference A36584
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A:Reference number: S00864; MUID:88166645; PMID:3327687
A:Accession: S00864
A:Molecule type: mRNA
A:Residues: 1-755, 'U', 757-1163 <CO3>
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830
A:Note: part of this sequence was confirmed by protein sequencing
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on m
C:Genetics:
A:Gene: GDB:ITGAX; CD11C
A:Cross-references: GDB:119758; OMIM:151510
A:Map position: 16p11.2-16p11.2
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>
F:20-1107/Domain: extracellular #status predicted <EXT>
F:149-319/Domain: von Willebrand factor type A repeat homology <WMA>
F:1108-1133/Domain: transmembrane #status predicted <TM>
F:1134-1163/Domain: intracellular #status predicted <INT>
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Db 141 LQGRPGQECIKG-----NVDLVFLFDGSMISLPDEFOKILDFMCKDVMVKLSNTSYQFAAV 196
QY 175 QYSEEFRIHTEFEFQNNPNSILVKPITQLLGRTHATGVRKVIRELLNITNGARKNAF 234
Db 197 QFSTSYKTEFDFSDYVKKWDPDALLKHVKMLLNTTFGAINVYVATEVFEELGAPDAT 256
QY 235 KILIVITDGKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNITLASKPPRD 294
Db 257 KVLIIITDGE--ATDSNGIDAAND-----IIRVIGIGKHFQYKESQETLHKFASKPASE 309
QY 295 HVFQNNFEALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITSNGLLSTVG 354
Db 310 FVKILDTFEKLDLFTLEKQKIVIEGTSKDLTSFNMELSSGSIADLSRGHVAVGAVG 369
QY 355 SYDWAGGVF--LYTSKESKSTINMTRVDSMDNDAYLGAAA--IILNRVQSLVLGARYQH 412
Db 370 AKDWAGGFLDKADLODDTTFIGNEPTTPEVRAGYLGTYVTWLPFSRQKTSLLASGAPRYQH 429
QY 413 IGLVAMFR--QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQT 470
Db 430 MGRVLLFQEPQGGHNSQVQTHGTQIGSYFGGELCGVDVDDQGETELLIGAPLFYGEQ 489
QY 471 RGQSVCPPLPRQARAWQDAV--LYGEOQPWGRFGAALTVLGDVNGDKLTVAGAP 528
Db 490 RGRVFYIY-----QRRQLGFEVSELQGDGPYPLGRFGEAITALTIDNGDLVDVAVGAP 544
QY 529 GEEDNCAVYLFHGTSGSLSPSHSQRISAGSKLSPLOYFGQSLSGGQDITMDGLVDLTV 588
Db 545 LEE--QNAVYIFNGRHG--GLSPOPSQRIEQTQVLSGIQWFGRIHGVKDLGGDLADVA 601
QY 589 GAQGHVLLRSQVLRVKALMEFNPREVARNVPECDNVV--KGKEAGEVRVCLHVOKSTR 647
Db 602 GABSQMLVLSRRPVDVMTVLSFPAEIPVHEVECSYSTSNKMGKGVNITICQI--KSLY 660
QY 648 DLREGOISWVYTDIALDSGRPHSRVAFNETKXSTRRTQVGLGLTQTCETLKLQLPNCI 707
Db 661 POP--QGRVANLYTYTLQDGHRTTRRGLFPGGRHLLRNIAVT--TSMSCDTPSFHPVCV 718
QY 708 EDPVSPVILRLNLSL---VGTPLS--AFGN-----LRPVLAEADAQLFTALPFPEKNCGN 757
Db 719 QDLISPINVLSNLSLWEEEGTPRQDAQKGDIPILRPSLHSETWEI-----PFENKCGE 773
QY 758 DNICQDDLSITTFMSLDCLVCGPREFNVTVRNDGEDSYRTQVTFEPFLDLSVRKYS 817
Db 774 DKCEANLRVSFSPARSRLRLTAFASLSVELSLSNLEDAVYVQDLHPPGLSPKVE 833
QY 818 TLQORSQRSLRACES--ASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCPELPEERLLSKAL---SCNVSSPIFKAGHSVALQMMENTLVNS 887
QY 876 LGNKLLKANVTSENN-----MPTNKTEFQLELPVKYAVYVMTVSHGVSTKYLNFATSEN 931
Db 888 WGSVSELHANVTNNEDSLEDNSATTI---IPILYPINILQDQEDSTLYVSFTPKGP 944
QY 932 TSRVMOHQVQV---SNLGRSLP--ISLVLVPLVRLNQTWDRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKHMYQVRLQPSIHDHNTPLTBAVGVQPPSEGIETHOWSVQMEPPV--PCHYED 1002
QY 987 -RLPSSHSD--FLAELURKAPVNCISIAVCQRIQCIDIPFFGIQEEFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAAEPCLPGALFRCPVW-----FRQEILVQVIGTLELVEITE 1044
QY 1044 TSHNHLIIVSTABILNDSVFTLLPGQAFVRSQETKVEPPEVFNPLPLIVGSSVGGLL 1103
Db 1045 AS--SMFSLCSSGISISNSKHFLYGSNASL--AQVVMKVDDVYEQMLYLYVLSGGIGLL 1102
QY 1104 LLALITAAALYKLFKKRQVKKMMSEG--GPPGAP 1136
Db 1103 LLLLIIVLYKGVFFRNILKERKEAGRVNGIP 1136

RESULT 5
I56126

lymphocyte fuction-associated molecule-1-alpha - mouse
C;Species: Mus musculus (house mouse)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56126
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.
J. Immunol. 147, 369-374, 1991
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit
A;Reference number: I56126; MUID:91268576; PMID:2051027
A;Accession: I56126
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1163 <RES>
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:G198785; PIDN:AAA39426.1; PID:G19878
C;Genetics:
A;Gene: LFA-1
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;151-315/Domain: von Willebrand factor type A repeat homology <VMA1>

Query Match 26.1%; Score 1533.5; DB 2; Length 1163;
Best Local Similarity 34.0%; Pred. No. 4.6e-97;
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;
QY 1 FNLDTENAMTFOENA--RGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGLSYQCDYSTGCEP 59
Db 24 YNLDTPTQSFQAQGRHFGYQLQIEDG--VVGAPGE---GDNTGGLVHCRTSSFECP 79
QY 60 IRLQVPEAVNMSLGLSLAATTSPQLACGPTVHTCSENTYVVKGLCFLFGSNLRQPPQ 119
Db 80 VSLH--GSNHTSKYLGMTLATDAAKGSLACDPLGSLTCDQNTVLSGLCYLPQSLGPGML 138
QY 120 KPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKEFVSTVMEQLKKSKTLFSLMOYSEE 179
Db 139 QNRPAYQECMKGVLDVFLFDGSQLDRKDFKILEFMKVMRKLSTNTSYQFAAVQFSTD 198
QY 180 FRHFTFEKF--QNNPNRSLVVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKILI 238
Db 199 CRTEFTFLDYVKNKNPDLVLSGVQPMFLTTNTFRAINVVAHVFFKEGSGARDATKVLV 258
QY 239 VITDGSKFGDPLGYEDVPEADREG-----VIRYVIGVGDAFRSEKSRQELNITIASKP 291
Db 259 IITDG-----EASDKGNISAAHDITRYIIGIKHFVSVQKQTLHIFASEP 304
QY 292 PRDHVQVNNFALKTIQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAAITSNGLPLS 351
Db 305 VBEFVKILDTFEKLDLFTDLQRIYAIIEGTNRQDLTSFNMELSSGSIADLSRGHVAVG 364
QY 352 TVGSDVWAGGVF--LYTSKESKSTINMTRVDSMDNDAYLGAAA--IILNRVQSLVLGAPR 409
Db 365 AVGAKDWAGGFLDLREDLQCATVCGEPLTSDVRCGYLGYTVVWMTSRSRPILLAGAPR 424
QY 410 YOHIGLVAMFR--QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467
Db 425 YOHVGQVLLFQAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGEAELLIGAPLFF 484
QY 468 EQTRGQSVSVCLPRQARAWQDAVLYGEOQPWGRFGAALTVLGDVNGDKLTVAGIA 527
Db 485 GBORGGRVETY---QRRQLFEMVSELOQDDPGYPLGRFGAATALTIDNGRLTIDVAVGA 541
QY 528 PGEENRGAAYLPHGTSGSGISPSHSQRISAGSKLSPLOYFGQSLSGGQDITMDGLVDLT 587
Db 542 PLEE--QNAVYIFNGKPG--GLSPQPSQRIQGAQVFGIRWFGRSIHGVKDLGGDLRADVV 598
QY 588 VQAQGHVLLRSQVLRVKALMEFNPREVARNVPECDNVQVVKGEAG--EVRVCLHVQKST 646
Db 599 VGAERGVVVLSSRPVVDVVTLSFSPPEIPVHEVECSYSAAREEQKHGVKLKACFRKPLT 658
QY 647 RDRLEGQIQSVVTVYDLDLDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNC 706
Db 659 PQ--FQGRLLANLSYTLQDGHMRMSRGLFPDPSHELSTNTSITP--DKSCLDFHFFHTIC 715
QY 707 IEDPVSPVILRLNLSL---GTPLSAFGN--LRPVLAEADAQLFTALPFPEKNCNDNICQ 762
Db 716 IQDLISPINVLSNLSLEEGTPRQKGRAMQPIILRPSIHTV--TKEIPFKNKCGEDKKCE 774

763	QY	DDL\$ITFSFMSLDCLVVGGP-----REFNVTVTRNDGSDSYRTQVTFPPDLDLSRVKV	816
		: : : :	
775	Db	ANLTGSPARS-----GLRLMSSASLAVETWLSNGSDAYWVRLLDDFPGRGLSPRKV	827
		: : : :	
817	QY	STLQNQRSQRWLACBSASTEVSGAL-KSTSCSINHPIPFENSEYFNITFDVDKSKAS	875
		: : : :	
828	Db	EMIQ---PHSRMPVSCBEL--TEGSULLTKLKCNVSSPIFKAGQSVLQWMFTNLNSS	889
		: : : :	
876	QY	LGNKLILKANVTSEN-NMPTNKTFFOLEPLVKYAVVMVTSHGVSATKYLNFTASENTSR	934
		: : : :	
883	Db	WEDFVELNGVHCENENSLOEDNSAATHIPVLYPVNILKEQENSTLYISFTEKGPKTQ	942
		: : : :	
935	QY	VMOHQYQVSNLQORSLPISIVELVPVRLNOTVIWDREQ----VFESNNLAS-----TCHTK	986
		: : : :	
943	Db	QVQHVGQV----RIQP\$AYDHNMP-T-LEALGVGPRPHSEDLIITYTWSVQTDPLVTCHE	996
		: : : :	
987	QY	E-RLPHSHDFLAELRKAPVNC\$IAVCORIOCDIPFGIOEENFATLKGNSLPDWIKTS	1045
		: : : :	
997	Db	DUKRPSSE---AEQPCUPGV-----QRFCLPV---RWELLIQVTGVSELKEIKAS	1042
		: : : :	
1046	QY	HNHILLIYSTAEILFNDSVFTLLPGQGA\$FVASQTKEYPEFVNPFLPIVGVSSVGGLLLL	1105
		: : : :	
1043	Db	-SITLSCSL\$SVSNSKKHFLYGSKA-SQAQLVAKVDLTHEKMLHVYVL\$IGGLGVLL	1100
		: : : :	
1106	QY	ALTAAIYKLGFFKRYQKDM-M-SEG\$PPGABP	1136
		: : : :	
1101	Db	FLIFLAIYKVGFFKRNLRKEAMEADGVNGSP	1132
		: : : :	

RESULT 6
A53213
integrin alpha-E chain - human
C/Species: Homo sapiens (man)
C/Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C/Accession: A53213
J/Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A>Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit. U
A/Reference number: A53213; MUID: 941164962; PMID: 81119947

Query Match	19.6%	Score 1149;	DB 2;	Length 1179;
Best Local Similarity	29.1%	Pred. No. 1.5e-70;		
Matches 343;	Conservative 211;	Mismatches 451;	Indels 172;	Gaps 39;
QY	45	GSLVQCDYSTGS--CEPI-RLQVP-----VEAYNMSLGLSLAATTSPQQLACGPTVHQ	95	
DB	65	GPHRCSLVQDEILCHVEHVPIFKGHRGVTVVRSHGVLCI-----QVLVRP--HS	117	
QY	96	TCSENTVYVGLCFPLGSLNRQQPQ-----	119	
DB	118	LSSELT---GTCSLLGPDLPQAQNFLENLLDPDARVDTGDCYSNKEGGEDVNTA	174	
QY	120	KFPALRGCPQED-----SDIAPLIDGSGSIIPHDPRRMKPEFVSTWQEL--	164	
DB	175	QRRALEKEEEDKEEEDDEEAEAGTETAILDGGSIDPPDQRAKDFISNMNRFEY	234	
QY	165	KSKTILSLMOYSEFRIHTFKFQNNPNRSLVKPITQLLGRTHATGVKVIKRELLN	224	
DB	235	KCFECNFALVGGYGVITQFEFLDRSDQVMASLARVQITGVSTVKTASAMQHLDSIFT	294	
QY	225	ITNGARKNAFKILIVITDGEKFGDPLGVEDYIPADREGVIRYVIGVDAPRKSROEL	284	

	Db	295	SSHSGRKASKVMVLLTDCGIFEDPLNLITTVINSPKMQGVRFPAICVGEEFKSARTAREL	354
	Qy	285	NTIASPPRDHVPQNNFALKTIQNLREKIPAIETQTGSSSPSEHEMSQBGSAAIT	344
	Db	355	NLIASDPDETHAFKVNYMALDGLLSKLRYNIISMEGT--VGDALHYQLAQIGFSAQIL	411
	Qy	345	SNGP-LLSTVSGVDWAGGWFLY-TSKEKSTFFINMTRVDSMDNDA----VLGYAAAAILRN	398
	Db	412	DERQVLLGAVGAPDWGSGALLYDTSRRRGFRFLNQTTAAADAADAAQAQSYLVLYGAVAVLHK	471
	Qy	399	RVSQSLVLAGPRYOHIGLVAWER-QNTGMWESNANV-KGTQIGAYFGASLCSDVDVDSNGST	456
	Db	472	CSUSYVAGAPQYKHGG-AVFELQKGREARASFPLVEGEOMGSIYFSELCPVDIDMDGST	529
	Qy	457	DLVLIGAPHYEQTRGCQSVCLPRGQRARWCDAVLXGEQQCPQREGAALTVLGDVN	516
	Db	530	DFLVLAAPPFHVHGEGRVVVYRLSE-QDGFSLARILSGHPGFTNARFGFAAMAGDLS	588
	Qy	517	GDKLTVAIGAP-----GEDNR--GAVYLPHGTSGSGISPSHQSORTAGSKLSRLOQYFGQ	570
	Db	589	QDKLTVAIGAPLEGFCAGDDGASFGSVIYNG-HWDGLSASPQRSIRASTVAEPLQLOYFGM	647
	Qy	571	SLSGGODLTWGDVLDTVGAQGHVLLLRQPVLVRKAIMEFNPREVARNVFECDNDQVKVG	630
	Db	648	SWAGGFDIQDGLADIIVTGLQAVVFRSPVRLVKVMAFTSALP-----IGF	697
	Qy	631	KEAGEVRVCLHVOKSTRDLREQIQSVVTYDIALDSGRPHSRVAFNETKNSTRRTQVIL	690
	Db	698	NGVVNVLRCFEI-SSVTASESGUREALLNFTLDVDVGQRRLQCSDVRSCLGLREWS	756
	Qy	691	GLTOTCETLKLOLPN---CIEDPVGPVILRLNFSLVGTPLSFAGNLRPVLAEADAORLETF	746
	Db	757	SGSQLEDL-LLMPTBEGELCEEDCFNASVKVSYQL-QTPGEGQTDHPQPIILDTRYTEPFPI	814
	Qy	747	ALFPPEKNCNDMIQDDDISITSPFMSLDCLVVGGRPEFNVTYVRNDCGDSXYRTQVTFP	806
	Db	815	FOLPYEKACKNKLFCAVELQLA-TTVSQOQLVGLTKELTNLINTLNSGDSYMTSMALN	873
	Qy	807	FPLDLSYRKVSTLQONORSQRWRWLACESASTEVSGALKSTCSINHPIEPENSEVTFNI	866
	Db	874	YPRNLQ-----LKRQKPSPNIOCDDPOPV--ASVLIMNCRIGHPVL-KRSSAHVSU	923
	Qy	867	TFPDVDSKASLGKLLKANVTSENN----MPRTNKTEFQ--LELPVKYAVYMYVVTSHGV	919
	Db	924	VWOLEENAPERNRADITVTVTNSERRSLANETHTLQFRHGFVAVLSKPSIMTYVNTQOGL	983
	Qy	920	S--TKYINFITASENTSRVMQHQOVSNLGORSHIPSLVFLVPVRLNQTVIDRPQVTFSE	977
	Db	984	SHKBEFLFVHGEN--LFGAEYO-----LQICVPTKRLGLOVAAVKKLRTQJ	1028
	Qy	978	NLSSTCHTERLPSHSDFLAEALRKAPWNCSIAVCQRIQCDIPFFGIQEIBFNATLKGMLS	1037
	Db	1029	ASTVCTWSQERACAYS-VQHVZEWHISVCVIA-----SDKENVTVAEIS	1073
	Qy	1038	FDWYIKTSHNHLIVST-----AEILFNDVSFTLLPGQGAFVRSQOTETKVEEFVENPL	1091
	Db	1074	WD-----HSEELLKDVTELQILGEISFNKSLYEGLNAENH-RTKITVTVFLDXDKYHSL	1125
	Qy	1092	PLIVGSVGLLLLLALITAALYKGLKFFKROYKOMMGE	1128
	Db	1126	PILIIGSVGLLVILVILVLFKCGFFKRYQOOLNLE	1162

RESULT 7
A45226
integrin
C;Species
C;Date:
C;Access
R;Briefer
J. Biol.
A.Ti-le:

A;Reference number: A45226; MUID:93155124; PMID:8428973

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151 <BRI>

A;Cross-references: UNIPROT:P56199

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: sequence extracted from NCBI backbone (NCBIP:124326)

F;142-317/Domain: von Willebrand factor type A repeat homology <WVAL>

```
Query Match      18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.4e-67;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

QY 1 FNLDTENAMTFOENARG-FCQSVVQL---QSGRVVVGAPQEIIVAANQSGSLYQCDISTGS 56
Db 1 FNVVKNMTFSFVEDMFGYTVQYENBEGKVLGSLVPGQPKNTGVDVYKCPVGRGE 60

QY 57 CEP-IRLOPVEA-----VNMSLGLSLAATTSPQLLAGCPTVHQTCSNTTVYKGL 106
Db SLPCVKLDLPVNTSIPNVTVENKMTFGSTL-VTPNPGGFLACGLYAVRCGLHLYTGI 119

QY 107 CFLFGSNLRQOPKPFALRGCPQESDIAFLDGGSIIPHDFRMKPEFVSTVMBOLK- 165
Db 120 CSDVSPTFQVNSIAP--VOECSTQ-LDIVIVLDGGSNIYPWD--SVTAFNLDLKRMDI 174

QY 166 -KSKTLFSLMOYSEEFRIHFTKFEFQNNPNRSLVVKPITOLLGR-THTATGVRKVIRELL 223
Db 175 GPKQTQGVIGVQENWTFENLKYSTETEVLAANKIVQGRQWMTALGTDRTARKEAF 234

QY 224 NITNGARKNAFKILIVITDGEKFGDPLGVBDVPEADREGVIRYVIGVDADR-SE 278
Db 235 TEARGARRGVKVMVITDGEH-DNHLKVKVQDCEDENIQFSAIILGSYNRGNLSTE 293

QY 279 KSKQELNTASKPRPHVFOVNNPEALKITQNLREKIPAEIQTGTGSSSPFHEMSQEG 338
Db 294 KFVEETKSIASEPTEKHFNVSDLAIVITKLGRIFALEATADQSAASFEMESQTG 353

QY 339 FSAAITSNGLPLSTGVSYDWAGGVFLVTSKE-----KSTF-INMTRVDSMDNDAYLGAA 392
Db 354 FSAHVSQDWMLGAVGAYDWNGIVVMQKASQIIIPRNTFNVESTKKNELP-ASYLGTV 412

QY 393 AILNRVQSL-VLGPARYOHIGLVAMFRONTGMESNANVKGTQICAYFGASLCSDVD 451
Db 413 NSATASSGDVLYTAGQPRYNTQVLIYRMEDGNIKILQTLSEQIGSYFGSLITLTDID 472

QY 452 SNGSTDVLVLGAPHY-----YEQTR-GGQVSVCLPRGQARWQCDVILY 495
Db 473 KDSNTDILLVGPAMYMGTEKEBQGVYVYALNQTRFEYQMSLEPIKQTCSSRQHNSCTT 532

QY 496 GBQGGQPWG-RFGAALTVLGDVNGDKLTDVAIGAPGBEDNRGAVLYLPHGTSGSGISPSHQ 554
Db 533 ENKNFCGARFGTAIAVKDLNLGDGNDIVIGAPLEDHGGAVIYTHG-SGKTIRKEYAQ 591

QY 555 RIAGSKSLRLOVFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRLSPVLRVKAIMEFNR 614
Db 592 RIPSQGDGKTLKFPFGS IHGEMDLNGDGLTDTVIGLGAALFWSRDAVAVKVTWVNFEN 651

QY 615 EVARNVFCNDQVVKKEAG--EVRVCLHVQ-KSTDRDLREGIOISVTVYDIALDSGRPH 671
Db 652 KVNIOKKNCH---MEKETVCINATVCFVKLSKEDTIVEADLQ----YRVLDSLRLQI 704

QY 672 SRAVFNET-----KNSTRQQTQVLGLTQTCETLKLQPLNCIEDPVPVIRLNSLVCT 725
Db 705 SRSFSGTQERKQVQIRNITRKSEB-----TKHSFYMLDKHDFQDSVR---ITLDFNLT-D 755

QY 726 PLISAFGNLRPVLAEDAQRLLTALFPPEKCNKNDNIQDDLSITFSPMSLDCLVVGGPRE- 784
Db 756 PENG-----EVLDDSLPNSVHEIYIPAKCGNKEKICISDLSHVATTEKDLIVRSQNDK 810

QY 785 FNVTVTVRNDGDSRYRTQVTFPPLDLSYRKVSTLQNRQSRSRWRLACASASTEVSGAL 844
Db 811 FNVSLTVKNTKDSAVNTRTIVHYSPLNVFSGIEAIQKD-----SCSEN----- 853
```

```
QY 845 KSTSCSINHPIFBNSEVTEFNITFDVDSKASLGN-KULLKANVTSENMMPTNKTEFOLE 903
Db 854 HNTCKVGYDFLRGEMVTFKILFQFNTSYLMENVTIYLSATSDSEBPPETLSDNVVIS 913

QY 904 LPVKYAVVMVVTSHGVSTKVLNFTASENTSRVMQHQYQVSN-----LGQRS-----L 950
Db 914 IPKYEVLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPMPEL 972

QY 951 PISLVP-----LVPVRLNQTVIWRDPQVTFSENLSSTCTHKE-----RLPS 991
Db 973 KLSISFPMNTSGVPVLYPTGLSS-----SENANCRPHIFEDPFSINSKQWTT 1021

QY 992 HSDFLAELKAPVNVGSIACVQRIQCDIPFGIOE-----EFNATLK 1033
Db 1022 STD---HLKRGTLDCNTCKFATITCNLTSSDISQVNVSLILWKPTFIKSYFSSNLITR 1078

QY 1034 GNLSFDWYIKTSNHLILIVSTABILEFNDSVFTLLPGQAFVRVRSQTEKVEPFEVNPPLP 1093
Db 1079 GEL-----RSENASLIVSSN-----OKRELAIQISKQGLPGRVPL 1114

QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKQYKDMMS 1128
Db 1115 WVILLSAFAGLLMLLILALWKIGFFKPLKKMK 1151

RESULT 8
I45914
Integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Ramada, I.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:I25886; NID:G439695; PIDN:AAB59255.1; PID:G439695
F;161-336/Domain: von Willebrand factor type A repeat homology <WMA2>
```

Query Match 18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 4.6e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

```
QY 1 FNLDTENAMTFO-ENARGFGQSVVQL---QSGRVVVGAPQEIIVAANQSGSLYQCD--DYST 54
Db 19 YNVGLPKALFSGPSSEQEGYAVQQFIPNKGWLLVGSFWSGFPKRMGDVYKCPVDLST 78

QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQLLAGCPTVHQTCSNTVYKGLC 107
Db 79 TTCEKLNLTSTSMGNNVTMKTNMSLGLTLTRNVGTGGFLTCGLPLWAQCGSQYTTTGV 138

QY 108 FLFGSNLRQOPKPFALRGCPQESDIAFLDGGSIIPHDFRMKPEFVSTVMEOLK-- 165
Db 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVVCDESNIYPWD--AVKNLEKFEVQGLDIG 194

QY 166 KSKTLFSLMOYSEEFRIHFTKFEFQNNPNRSLVVKPITOLL-----GRTHATGVRKVIRE 221
Db 195 PKTQOMGLIQYANPRVFNLTNFKSKD---EMIKATSTQTFQYGGDLTNTFKAIQIARDT 251

QY 222 LNIITNGARKNAFKILIVITDGEKFGDPLGVBDVPEADREGVIRYVIGV-----GDAPR 276
Db 252 AYSTAAGRGPGATKVMVVVTDGESH-DGSKLKAVIDQCCKNILRFGIAGVLYNRLALD 310

QY 277 SEKSRQELNTASKPRPHVFOVNNPEALKITQNLREKIPAEIQTGTGSSSPFHEMSQ 336
Db 311 TKNLKIKETKATASIPETERHFFNVSDADLLEKAGTIGEQIFSGIEGTVOQ-GDNFQWMSQ 369

QY 337 EGFSAAIT--SNGPLSLSTVGSYDWAGGVFLVTSKEKSTFINMT--RVSDSMN-DAYLGVA 391
Db 811 FNVSLTVKNTKDSAVNTRTIVHYSPLNVFSGIEAIQKD-----SCSEN----- 853
```

Db 370 VGEAEVSPQNNILMLGAVGAWDMSGVVQKTPHGHLIFSKQAFQEQILQDRNHSSYLGS 429
Qy 392 AAILNRVOSLVGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYFGASL 445
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNVTIQSQRGDQIGSYFGSVL 484
Qy 446 CSVDVDSNGSTDVLVIGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLGEQGPWG 503
Db 485 CAVDVNKDTITDVLVAGAPMYMDLKKBEGRVYLFTITKG--ILNWH--QFLEGNGLENA 541
Qy 504 RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLPHGTSGSGISPSHSQRLAGS--KL 561
Db 542 RFGSAIALSDINNDGFDNDIVGSPLENQNSGAVIYNGHEGM--IRLRYSQKILGSDRAF 600
Qy 562 SPLQYFGSLSGGQDLTMDGLVDLTVGAQGHVLLRLSQPVLRVKAIMENPREVARNVF 621
Db 601 SSHLQYFGRSLDGYGLDNGSDITDVSVGAFOGVQVQLMSQIADVSVDASETPKKI--TL 658
Qy 622 ECNDQVYKGEAGVRVCLHVQKSTRDLRREGQIQSVVTVYDLALD---SGRPHSRAVEN 677
Db 659 NKNRAE-----KKLCLF-----SAKFRPTNNQNNQVAIVNITIDEQFSSRVLSRGLFK 707
Qy 678 ETKNSTRRQTVLGLTQCE--TLKLQLPNCIEDPVSPIVLRNLFSL--VGTPLSAFENL 733
Db 708 ENNERCLOKTMIVSQACRCSEVITHIQEPS---DIISPLNLCMNISLENPCGT----- 756
Qy 734 RPLVAEDAORLFTALFPFKKCGNDNICQDDLSITF-----SFMGLDCLVGGPREFNVTV 789
Db 757 NPALAEAYSETVKVFSIPFKDCGDDGVCISDLVLNVQOLPATQOQPFIVSNQNKRLTFSV 816
Qy 790 TVRNDGDSVYQTQVTFEPFLDLSYRKVSTLQNRQSRQSWRLACASST--EVSQALKSTS 848
Db 817 QLNKKESAVNTEIVDFSENLFF-----ASWSMPVDGTEVTCQIASQSKSVT 864
Qy 849 CSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTFFOLELPVKY 908
Db 865 CNVGYPAKSKQVTFITNFDNLQ--NLQNASISFRALSQSENNMADNSVNLKLSILY 923
Qy 909 AVYMWTVSHGVSTKYLVNFTASENTRVMQHOYQVSNLQOR-----SLPTSLFLV 958
Db 924 DAEIHT--RSTNINFEYSLDGNVSVV--HSFE--DIGPKFIFSIVTGVSPVMSA--- 976
Qy 959 PVRLNQTVIMDRPQVTSSEN--LSSTCHTKB-----RLPSHSDFLAE- 998
Db 977 -----SVIIHPQVTKDNPLMLVLTGVHTQAGDISCEABINPLKGTQSSVSFKSEN 1030
Qy 999 LRKAPVNCSTAVCQRIQCDIPFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABI- 1057
Db 1031 FRHIKELNCRATASCNIMCWLRLDLQVKGVEFLNVSTRINWGTFAASTFQVLTAAABID 1090
Qy 1058 LFNDVSFTL-----LPCQGAFVRSQETKVEPPE--VPNPLPLIVGSSVGLLILALITA 1110
Db 1091 TYNPQIVYIEBNTWIP-----LTIMKPKHEKVEVPTGVIVGVSIAGLILLLALVA 1140
Qy 1111 ALYKLGFPKROYKDM 1125
Db 1141 ILWLKGFPRKRYKYM 1155

RESULT 9

S44142
VLA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S44142
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Danjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PID:CAA82877.1; PID:g4747
F;169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1072; DB 2; Length 1178;
Best Local Similarity 28.1%; Pred. No. 3.1e-65;
Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

Qy 1 ENLDTENAMTFQ--ENARFGQSVVOL--OGSRVWVGAPQEIIVAANORGSLYQC--DYST 54
Db 27 YNVLPGAKIFSPSESEQGYVQOQTNPGNMLLVGSPWSPGPNRMGDVYKCPVDLPT 86
Qy 55 GSCEPIRLQ-----VPVEAVNMSGLSLAATTSPQLLACGPTVHOTCSENTYVKGCL 107
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCGFLWAHQCNGQYATGIC 146
Qy 108 FLFGSNLRQPOQ--RFPPEALRCGPQEDSDIAFLIDSGSIIIPHDPRMKEFVSTVMEQL 164
Db 147 ----SDVSPDFQFLTSPFAVQACPSL--VDVVVVVCDENSIYP--WEAVKNFLVFTVGL 199
Qy 165 K--KSKTLFSLMOYSEEFRIHFTKFEQNNPNRSLVKPITQLIG--RTHATGVRKVVIRE 221
Db 200 DIGPKTKQVALIQVANEPRILFNLDPETKEDMVQATSETFQHGDLTNTFRALEFARDY 259
Qy 222 LLNITNGARKNAFKILIVITDGEKFGDPLGYEDVPEADREGVIRYVIGV-----GDAFR 276
Db 260 AYSQTSGRPGATKVMVVTVDGESH--DGSXLKTVIOQCNDDEILRFGI AVLGYLNRNALD 318
Qy 277 SEKSRQBLNTIASKPRDRHVQVNNFALKTIQNLBEKTFARIGTQSGSSSEHEMSQ 336
Db 319 TKNLIKEIKATISTERYFNFVADEAALLEKAGTLGEQIFSGIEGTVQG--GDNFQMEWAQ 377
Qy 337 EGPSA--AITSNGPLLSVGSYDWAGVFLYTSKEKSTFINMT--RVSDSMN--DAYLGYA 391
Db 378 VGSADYAPQNDILMLGAVGAFDWSGLTVQETSHKPVIFPKQAFDQVLQDRNHSSFLGS 437
Qy 392 AAILNRVOSLVGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYFGASLCS 447
Db 438 VAAISTEDGVHFAVAGAPRANVTGQIVLYSVN---QGNVTVIQSHRGDQIGSYFGSVLCS 494
Qy 448 VDVSNGSTDLVLIGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLGEQGPWGFR 505
Db 495 VDVDDKDTITDVLVAGAPTYMNDLKKERKGVLFITIKILNQHO---FLEGFTGTGNARF 551
Qy 506 GAALTVLGDVNGDKLTDVAIGAPGERDNRGAVLPHGTSGSGISPSHSQRIAGSKLSPR- 564
Db 552 GSAITAAALSDINNDGFDNDIVGSPVENENSGAVIYNGHQGT--IRTKYSQKILGNGAFPR 610
Qy 565 --LQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRLSQPVLRVKAIMEFNPREVARNVPEC 623
Db 611 HLQFFGRSLDGYGLDNGSDITDVSIGALGVQLWSQSIAADVAIEALFTP----- 660
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVTVYDLALD---SGRPHSRAVENET 679
Db 661 -DKITLANKDAKITKLKCFRAEFPAGQNNQV--AILFNMTLDADGHSRVTSGVFRFN 717
Qy 680 KNSTRROTQVLGLTQTCET--LKLQLPNCIEDPVSPIVLRNLSFVLGTPLSAFGNLRPVL 737
Db 718 SERFLQKNMVNVEVQKCSHHISIQKPS---DVVNPLDLRVDLSLENPGTS-----PAL 768
Qy 738 AEDAQRFLTALFPPEKNGCNDNICQDDLSI-----TFSPMSLCLVVGGRFENVT 788
Db 769 EAYSETVKVFSIPFYKECGSDGICITSDILVDQOLPAIQTQSF-----IVSNQNKRLTFS 823
Qy 789 VTVNDGDSVYQTQVTFEPFLDLSYRKVSTLQNRQSRQSWRLACASST--EVSQALKST 847
Db 824 VILKNRGESANTVVLAEFSENLF-----ASFMPVDGTEVTEVGSQSKSV 871
Qy 848 SCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTKE--FOLELP 905
Db 872 TCDVGYPAKSKQVTFITNFDNLQ--NLQNASINFAQFSESQ--ETNKADNSVSLATP 928
Qy 906 VKYAVYMWVTVSHGVSTKYLVNFTASENTRVMQHOYQVSNLQOR-----SLPTSLV 955

Db 929 LLYDAELHILT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSM 984
QY 956 FLV-----PVLNQTIVDRPQVTF-SENLS 980
Db 985 LVTHIHPQYTKENKPLLYLTGTQDQAGDISCTAENPLKPHTA-----PSVSFKNER 1040
QY 981 STCHTERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFFGIGOEENATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMKAEYFINVTRVWRT 1080
QY 1041 YKTSNHLIYSTAEILFNSVFTLLPGGAFVRSQETKYEPFVFNPLPLIVGSSVG 1100
Db 1081 FFASTFQTQVLTAAABIDTHNPOLPFIENAVTIPLMIMKTEKAEVPT--GVIIGSIIA 1138
QY 1101 GILLALLITAALYKLGFEKQYKDM 1125
Db 1139 GILLALLAMTAGLWKGFEKQYKDM 1163
RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907
A;Note: the authors translated the codon GAR for residue 802 as Gln, GTC for residue 803
R;Catmell, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:I24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VMA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 5e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;
QY 1 ENLDTENAMTFQ-ENARGFQSVVOL---QGSRVVVGAPQEIIVAAQNRGSLYQC--DYST 54
Db 30 YNGLPEAKIFSGSSEQYVQVQVQPINPKGNWLLVGSFWSGPPENRMGDVYKCPVDLST 89
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVKGLC 107

Db 90 ATCEKLNLTSTSIPIVNTKMTNMSLGLILTRNMGTGGELTCGLPLWAQCCNQYITGVC 149
QY 108 FIFGSLNRQOPQKPPALRGCPQEDSDIAFLIDGSGIIPHDPFRMKCFVSTVMBOLK-- 165
Db 150 SDISPDPF-QLSASFSPATQPCPSL-IDVVVVCDESISYPWD--AVKNFLEKFEVGLDIG 205
QY 166 KSKTLFSLMOYSEEPRIHETFEFQNNPNRSLVKPIITQLLG-RTHATATVRKVIRELN 224
Db 206 PKTQVGLLIQYANPRVFNLTNYTKKEEMVATSTQSYGDLTNTFGAIQVARYAYS 265
QY 225 ITNGARKNAFKILIVITDGEKFGDPLGYEDVIPEADREGVIRVIGV-----GDAFSEK 279
Db 266 AASGGRSATAKVVVVVTDGESH-DGSMKXAVIDOCNHDNLIRGLIAVLGVLNRLDTKN 324
QY 280 SRQELNTIASKPPRDHVQVNNFEALKTIONOLREKIFAIEGTQTGSSSFHEHMSQEGF 339
Db 325 LIKEIKAIASIPTEYFFNVNDEAALLERKAGTLGEQIFSTIEGTVOG-GDNFQMEMSQVGF 383
QY 340 SAAITSNRP--LLSTVGSYDMAGVFLYTSKEKSTFINMT--RVDSMDN-DAYLGYAAAI 394
Db 384 SADYSSQNDILMGAAGAFGWSGTIVQKTSHGHLIPPKOAFDQILODRNHSSYLGSVAA 443
QY 395 ILNRNRVQSLVLAGAPRYQHIGLVAMFRQNTCMWESNANV-----KGTQIGAYFGASLCV 448
Db 444 ISTGESTHFVAGAPRANYTGQIVLSVN-----ENGNIIVIQAHRGDQIGSFVGLCSV 498
QY 449 DVDSNGSTDLVLIGAPHYIEQTR--GGQVSVCPPLRGQARWQCDVAVLGEQQPMGRFG 506
Db 499 DVDKOTITDVLVAGAPMNSDLKKEGRVYLFETIKKGILQHQ--FLEGPEGIENTRFG 555
QY 507 AALTVLGDVNGKLTDAITVAGCEEDNRGAVLFGHTSGSGISPSHSORTAGS--KLSPR 564
Db 556 SAIAALSDILMDGFNDVIVGSPLENQNSGAVIYNGHQST-IRTKYSQILGSDGAFRSH 614
QY 565 LOYFGQSLGGQDLTMDGLVLDITVGAQGHVLLLRSQPLVRVRAIMEFNPREVARNVFEEN 624
Db 615 LOYFGRLSDYGLDNGDSITDVISGAFQVQVQLWSQSIADVAIEASFTEKI--TLVNKN 672
QY 625 DQVVKGEAGEVVRVCLVHVKQSTRDLRLEGQISQVVTYDLALD-----SGRPHSAVFNETK 680
Db 673 AQII-----LKLCF---SAKFRPTKQNNQVAIVYNTILDADGFSRVRTSRGLFKENN 721
QY 681 NSTRQTVLGLTQTC--ETLKLQLPNCIEDPSPVILRLNLSLVGTPLSAFGLRPLVLA 738
Db 722 ERCLQNMVNVNAQSCPEHIITQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALFPFPEKNCNDNICQDDLSITF---SFMSLDCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEGDGLCISDLVLDVRQIPAAQEQPFIIVNQNKRLTFSVTLKNK 832
QY 795 GEDSVTRQVTFPPDLDSYRKVSTLQNRQSRWRLACESAST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSVACDVGY 880
QY 854 PIFPENSEVTNTITFDVDSKASLGNKLLKANVTSNNMNPRTNKTEFOLELPVKYAVMV 913
Db 881 PALKREQQVTFINFDNLQ-NLQNASLSFOALSEQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSHGVSATKYNFTASENTSRVMQHYQVSNLQOR-----SLPISLFLV----- 958
Db 939 ---HLTRSTNINFISSDGNVPSIVHSFEDVGPKFIPSLKVTGTSVPVSMATVHIHQ 995
QY 959 -----PVLNQTIVWDRPQVTF-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLKIGT-----SSVSFSPKSNFR---HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFFGIGOEENATLKNLSFDFWIKTSHNH 1048
Db 1048 -----LNCRTASCNSVTCLWKDVKHMKGEYFVNVTVTRIWNGTFFASTFTOT 1091
QY 1049 LLIVSTAETI-LFNDSVFTLLPGGAFVRSQETTKVPEFVPNP-----LP--LIVGSSVG 1100

Db 1092 VOLTAAAEINTNPETIIVV-----EDNTVTIPLIMTKPDEKAEVPTGVILGSIIA 1141

QY 1101 GLLLLALITAAALKYKGFKKQYKDM 1125

Db 1142 GILLLLALVAILWKLGFFKKRYEKM 1166

RESULT 11

A35854

Integrin alpha-1 chain precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 09-Jul-2004
C;Accession: A35854; S11243

R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990

A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A
A;Reference number: A35854; MUID:90338125; PMID:2380249

A;Accession: A35854

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1180 <IGN>

A;Cross-references: UNIPROT:P18614; GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494

C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

F;170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.08; Score 1060; DB 2; Length 1180;
Best Local Similarity 27.48; Pred. No. 2.1e-64; Mismatches 478; Indels 234; Gaps 47;
Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

QY 1 FNLDTENAMTFOENARG-FGQSVVLQ---QGSRYVVGAPQEIIVAAVNRGSLYQCDYSTGS 56

Db 29 FNVVDKMSFSGPVEDMFGVTVQYENEBKWLIGSLVPLVGQPKARTGDVYKCPVGRER 88

QY 57 CEP-IRLQVPEA-----VNMSIGLSLAATSPQQLACQPTVHQTCSENYYKGL 106

Db 89 AMPCKVLDLPVNTSIPNVTEIKENNTFGSTL-VTPNPGGFLACGPLYAYRCGHLHYTTGI 147

QY 107 CFLFGSLNRQOPQKFPPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLK- 165

Db 148 CSDVSPFTQVNSFAP--VOECSTQ-LDIIVLDGNSIYP--WESVIAFLNDLKKMDI 202

QY 166 -KSKTFLSLMQYSEFRHFTFKFQNNPNRPSIVKPTITQLLG-RHTATGTVKRVIRELL 223

Db 203 GPKQTQVGIQYGENVTHFNWLNKYSTEEVIAANKIGROGGLQWTALGIDTARKEAF 262

QY 224 NITNGARKNAFKILIVTDGKFGDPLGYEDVIEADREGVIRVIVGVGDAFR-----SE 278

Db 263 TEARGARRGVKKVMVITVDGESH-DNYRLKQVIQDCEDENIQRFSAIALGHYNRGNLSTE 321

QY 279 KSRQELNTIASKPPDRHVFQNNFEALKTIQNLREKIFALEGTQTCGSSSFHEMSQEG 338

Db 322 KFVEIKSIASEPTEKHFNFVSDDELAVTIIVKALGERIFALEATADQSAASFEMESQTG 381

QY 339 FSAAITNGPLSLTVSGYDAGGVFLYTSKEKSTFINMT--RVDSMDND---AYLGAAA 393

Db 382 FSAHYSQDWMLGAVGAYDNGVTVMQKAMQVPHNTTQTEPAKNEPLASLYGTVN 441

QY 394 IILNRVQSLVGLAPRQHIGLVAMFRQNTGWESNANVKGTQITGAFGLASLCSDVDN 453

Db 442 SATIPGDVLVITAGQPRNHTQGVVIYKMDGNINILQTLGGEQIGSYFGSVLTTIDDKD 501

QY 454 GSTDLVLIGAPHY-----YEOTR-GGOVSVCPPLPRGQARWCCDAVLYGE 497

Db 502 SYTDLLLVGAPWYMGTEKEEGKQVYVAVNQVTRFYQMSLEPIRQTCSSLSKNSCTKEN 561

QY 498 QGQPMWG-RFGAALTVLGDVNGDKLTVDAIGAIPGEDNRGAYVLPHGTSGGSIPEHSQRI 556

Db 562 KNEPCGARFGTAIAAVKDLNVDGDFNVVIGAPLEDHAGAVYIYHG-SGKTIREAYAQRI 620

QY 557 AGSKLSFRLQYFGQSLSGGDLTMDGLVDLTGVAGHVLILRSQPLRVKAIMFENPREV 616

Db 621 PSFGDGKTLKFFGSGSIHGEDMDLNGDGLTDVVTIGLGAALFWARDVAVKVTMNFENPKV 680

QY 617 ARNVFECNDQVVKGBAG--EVRVCLHQV-KSTRDRLEGQISQVVTYDLDLSDGRPHSR 673

Db 681 NIOKKNCR---VEGKETVCINATMCHVKLSKEDSIYEADLQ-----YRVLDSLRQISR 733

QY 674 AVFNET-----KNSTRRTQVILGLTQTCETKLQLPNCI-----EDPVSPIVRLR 718

Db 734 SFESGTQERKIQRNITVRESE-----CIRHSFYLMDKHDFQDSVRVTL 776

QY 719 NFSILVGTPLSAFAGNLRPVLAEADAQRLETFALFPPEKNGCNDNICODDLSTIFSPMSLCLV 778

Db 777 DFNLT-DPENG-----PVLDDALPNSVHEHIPFAKDCNKRKCISDLTJLVNSTTEKSLLI 830

QY 779 VGGPRE-FNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSQSRWRSLACESASS 837

Db 831 VKSQHDKFVSLIVKKGDSAYNTRVVOHSPNLIIFSIGIEIQXD-----SCSEN-- 880

QY 838 TEVSGALKSTCSINHIPIFPENSEVTNITFDVDSKASLGNKLL-LKANVTSENMPRTN 896

Db 881 -----QNITCRGVFPFLRAGETVTFPKIIFQFNTSHLSENAIHLSATSDSEBPLESLN 933

QY 897 KTFEQLPLPVKAV---YMWVTSHGVSST-----KYLNTASENTSRVMQHOYQVSNL 945

Db 934 DNEVNTSIPKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVTYTKR 991

QY 946 GQRSPLP---ISLVF-----LVPVRLNQTVIMD-----RP-----Q 972

Db 992 GHFMPPELQLSISFFNLTADGYPVLPYG-----WSSSDNVNCRPSRLEDPFGINSKK 1045

QY 973 VTFS-----ENLSSTCHTKERLPDSDFLAELRKAPVNCSTAVCQRIQCDI--- 1019

Db 1046 MTISKSEVLKRGITQDCSTC-----GVATITCSLLPSDLSQVNVSL 1088

QY 1020 ---PFGIQIEEF---NATLGNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGF 1073

Db 1089 LWKPTF-IRAHFSSNLNLTRELK-----SENSLTSSN----- 1123

QY 1074 VRSQTEKVPFFVEVPNPLP---IVGSGVGLLIALIATAALYKLGFFKRYKDMSE 1128

Db 1124 RKRELAIQISKDGLPGRVPLWVILLSAFAGLLMLLILALWKLGFFKRPPLKKWEK 1180

RESULT 12

A41131

Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N;Alternate names: integrin alpha-4

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 09-Jul-2004

C;Accession: A41131; S16742

R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.

J. Cell Biol. 115, 1149-1158, 1991

A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte

A;Reference number: A41131; MUID:92064645; PMID:1840602

A;Accession: A41131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1039 <NEU>

A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:951484; PIDN:CAA37316.1; PID:951484

C;Superfamily: integrin alpha-4 chain

C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.38; Score 663; DB 2; Length 1039;
Best Local Similarity 22.78; Pred. No. 4e-37;
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDTENAMTFO-ENARGFGQSVV-QLQSR--VVVGAQEIIVAAV-----QRGLYQCDY 52

Db 41 YNLDPENALLYQSGSGTLFGYSVVLVSHGSKRMILVIGAPTASWLSNVSASVWNPVGAIRCGI 100

QY 53 STG---SCEPIRLQVP-----VEAVNMSLGLSLAATTSP-POLLACG---PTVHQ 95

Db 101 RKNPNQTCQEQSGSPSGEGCGTKCLEERNQWLGVTLSPQSGENGSIIVTCHEWKNIFY 160

QY 96 TCSENTYVKGICFLFGSLNLRQQQPKFPPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMK 155

QY 645 STRDLREGQIQSVVTVLALDSGR-----PHSRAVENETKNSRRROTQVLGLTQTCETL 699
Db 507 CFSYKGEVPGYIVLFTNMSLDVNRKASPRPFYSSNGTSDVTIGIQSSREANCRT 566
QY 700 KLOLPNCIEDPVPISVILRLNFSLVGTPLS-----AFGNLRPVLAEQAQR-LFTALFFPEK 753
Db 567 QAFMRKDVRLITPIQIEAAVHLGPHVISKRSTEEFPPLQPILOQKKEKDIMKKTINFAR 626
QY 754 NCGNDNICQDLSIT--PSPM-----SLDCLVGGPREFNVTVTVNRNDEGDSYRQTQVTF 807
Db 627 FCAHEN--CSADLQWSAKIGLPHENKTYLAAGSMKTMMLNLSFNAGDDAYETTLRVKL 685
QY 808 PLDLSYRKVSTLQNRQSRWRACASSTEVSGALKSTSCSNHPIFPENSVTNIT 867
Db 686 PVGLYFKILELEK-----QINCE--VTDNSGVVQ-LDCSICYIYVHLSRIDISFL 735
QY 968 FDVDSKASLGNKLLKANVTSEN--NMPRTNKTEFOLELPKYAVYVMVTSHGVSSTKYL 925
Db 736 LDVSSLGRABEDLSITVHATCENEEEDNLKHSRVTVAIPLKYEVLTVHGFVNPTSFVY 795
QY 926 FTASENTRV-----MOHQVQSNLQORSIP-LSLVFLVPVRLNQTWIDRPOVTFSEN 979
Db 796 GSDNENPETCWBEKXNLTHVINTGSMAPNVSVEIMVNSFS-----PQTDKLFNI 848
QY 980 -----SSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIPFGIOBEFNATLK 1033
Db 849 LDVOTTGCEHFENYQ-----RVCALEQ-----QKSAMQTLK 880
QY 1034 GNLSFDWIKYIKTSHNHLIVSTAE-----ILFN-----DSVFTLLPGGAFVRSQTE 1079
Db 881 GIVREF--LSKTRKLLCYIKADPHCLNFCNFCMKGESKEASVHIQLEGRPS-ILEWDE 936
QY 1080 TKVEPFV-----PNP-----LPLVGSVSGLL 1103
Db 937 TSALKFIRATGPEPNRPRVIELNKDENVHVLLEGLHHPKRYFTVITSSLLGLI 996
QY 1104 LLALITAAVLKLGFFKQYKDMSE 1128
Db 997 VLLLSIYVMWKAGFFKQYKSILOE 1021

RESULT 14
I58409
integrin alpha-9 chain precursor - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I58409; A49459
R:Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A>Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A:Reference number: I58409; MUID:94119603; PMID:8290272
A:Accession: I58409
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1035 <RES>
A:Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BARA04984.1; PID:g533327
R:Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A>Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A:Reference number: A49459; MUID:94064789; PMID:8245132
A:Accession: A49459
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-1035 <PAL>
A:Cross-references: GB:I24158
C:Superfamily: integrin alpha-4 chain
C:Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 7.5e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;

QY 331 EHEMSOEGFSAATISNGPLLSITVSGYDWAGGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389
Db 190 EHSCQAGIAGFTTEELVVMGAPGSPYWGATIKVLNLTD-NYLLKLNDEVMNRRYTLG 248
QY 390 YA--AAIILNRNVQSLVLGAPRYOHIGLVAMFR--QNTGMWESNANVKGTOIGAYFOASL 445
Db 249 YAVTAGHFSHPSTIDVVGGAPQDKGIVIFRADRRSGTLIKIFQASGKMGSYFGSSL 308
QY 446 CSVDVDSNGSTDVLIGAPHYEYQTRGGQVSVCPLEPGQARWQCDVAVLGEQGPWGRF 505
Db 309 CAVDLNAGDGLSDL-LVGAPMSEIRDEGQVTVY--INRNGGALIE-QUALTGD-GAYNAHF 364
QY 506 GAALTVLVDGVNGDKLTDVAICAGEEDNRGAVLPHGTSGSGISPSHSORIASGKLSPL 565
Db 365 GESIASLDDLDNDGFPDVAIGAKEDDFAGAVIYHGDAG-GIVPOYSMKLSGQKINPVL 423
QY 566 QYFGQSLSGODLTMDGLVDLTGVA--QGHVLLRSOPVLVRVKAIMEFNPREFVARNVPEC 623
Db 424 RMFGQISGIGIDMGNGYPDVTVGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
QY 624 NDQVVKGEAGEVRV-CLHVQ-----KSTRDLREGQIQSVVTVLALDSGRPHRAV 675
Db 483 HD-----GQOPVNCNLNVTTCFSFHGHVPBEI---GLNYVLMAVAKKEGQMGRVY 531
QY 676 F--NETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNFSL----VGTPLS 728
Db 532 FVLGETMGQVTEKLQTYMEETCRHYVAVHKKRVQDVSPIVEAAYSEHVTVGEER 591
QY 729 AFGNLRPVL-----AEDAQRLFTALFPFKNCGNDNICQDLSITFSFM--SLD----CL 777
Db 592 ELPLLETVLRWKKGOKIAQKNQTV---FERNCRSED-CAADLQLQGLKLLSSMDKTYL 647
QY 778 VVGPREFNVTVTVNRNDEGDSYRQTQVTFPFLDLSYRKVSTLQNRQSRWRACESASS 837
Db 648 ALGAVKNISLNISISNLGDDAYDANVSFNVSRELFF-----INMOKEMGISCELLES 701
QY 838 TEVSGALKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNK 897
Db 702 DFL-----KCSVGFPPMRSKSYEFVIFDTHLS--GEEVLSFIVTAQSG--NTER 750
QY 898 TE-----FOLELPKYAVYVMVTS-----HGVSTKYNFTASENTS---RVMOHQY 940
Db 751 SESLHDNTLVLMVPLMEHVDTSITGIMSPTSFVYGESVDAANFTQLDDLECHFPINITL 810
QY 941 QVSNLQORSIPISLVFL-VPVRLN-----QTVIWDPRQVTVSENLSSTCHTKER 988
Db 811 QVYNTGPESTLPGSSVSISFNNRSLSSGGAEMFHVQEMVVGQEKNCFSQKNPTPCIIPQEQ 870
QY 989 LPSHSDFLAELRKA-----PVCNCSIAVCQRIQCDIPFGIOBEFNATLKNLSFD 1039
Db 871 ENIFHTIFAFYTKSGRKVLDCEKPGISCLTAHCN-----FSALAKEESRTI-----D 917
QY 1040 WYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTEKVEP----- 1084
Db 918 IY-----MLLNT-BILKDDSSVIQ-----FMSRAKVKVDPALRVVIEAHGNPEEV 962
QY 1085 ---FEVFN---PPLPIVQ-----SSVCGLLLLALITAAVLKLGFFKQYKDM 1126
Db 963 TVVFEALHNLPRGYVVGWVIAISLLVGLIFLLAVLLMKMGFFRRYKEII 1015

RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HEP>

A:Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;

Best Local Similarity 24.8%; Pred. No. 8.8e-34;

Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

```
QY 336 QRGFSAAITSNGP--LLSTVGSYDMAGGVFLYTSKEKSTFINNTRVDS-----DMNDAYLG 389
Db 181 QAGFGIIFSDNSALVAGAPGSYVLOGQIYVQSLNRSV-VQATQESNTGTYSFDSYRG 239

QY 390 YAAAI--ILNRVQSILVLAGAPRYOHI-GLVAMPRQNTGMWESNANYKGTQIGAYFGASLC 446
Db 240 YSLALGDFNGGVQDYVVGTPRAESLWGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

QY 447 SYDVDSNGSTDVLIGAPHYEQTRGQSVSVCLPRGQARMOCDAVLYGEQ-----498
Db 297 VVDI--NNDTYDOLLVGA PMYMDGPAIQ-----RWEAGAVVYVLQNPVGVGPGA 343

QY 499 -----GPGWRFGCAALTVLGDYNGDKLTDVAIGAGEEDNRGAVLFGHGTSGS 546
Db 344 SNRLSUSLYLIGQIKRSRFLGSLTASIGDSNQDFNDVAIGAPYEGDAGAVYIYG-SAN 402

QY 547 GISPSHSQRIAGSKLS-PRLOQFPGSLGGQDLTMDGLVDLTVGAQ--GHVLLRSQPVL 603
Db 403 GLKSTPAQVLTPSTLGHSGITTFGSLQGGQMDKKNYPDLLVGAESANTAVLIRTPVV 462

QY 604 RKVAMEFNPREVARNVFECNDQVVGKEAGEVRVCLHVQKSTRDRLRREGQIQSV-----558
Db 463 SLDATLNTF-----IGINLENKTYE-LADGTMTVTSFIAMT 497

QY 659 -----VTYDLALDSG-RPHSRVAFNETKNSRRQTVGLL-TQTCETLKLQ 702
Db 498 CTTYTGNYLPDHIDISYITVVDSGIITANRRAMEFVDMDSMEITKRRRLAVSTQPCDELRAY 557

QY 703 LPNCIEDPVSPVLRNLFSLVGTPLSAFGN-----LRPVLAEADAQRLEFALPFPFEK 753
Db 558 VGNSEDKLTPIKVTIQLYDL-----NNDESRLQPHLELPIDMATMSTQTKQVSIQN 609

QY 754 NCGNDNICODLSITSFMSLCLVGGPREFNVTVTVRNDGEDSYRTQVTFPFDLSY 813
Db 610 NCVN-NICIPDLDTVTVT-PNLENVIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLQLQ 667

QY 814 RKVSTLQNRQSRWRLACESASSTEVSGALKSTSCSINHPIPPEN-----SEVTFNIT 867
Db 668 VRI-----ERRANMDFSVTCSUSD-----LRITCDTGNPMVGKNILFGLTLSTFQVS 717

QY 868 FDVDS-----KASLGNKLLKANVTSENMPRTNKTEFQLELPKYAVYVMVVTSHGVSTK 922
Db 718 GDKDSIEFYFKAESENS--EDPNTLENELNMVTPVTVDCITLKLSSAYPEIWMYSTQED 775

QY 923 YL-----NFTASNTSRVMQHOYQV-----SNLQORSLPIS-----LVFLVPV 960
Db 776 YVVPFPFPAKNASPADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKNEDGEYLFYLLGI 834

QY 961 RUNQTVIMDRPO-----VTFSENLSTCHTKELPSHSDFLAELRKAPV 1005
Db 835 MTEEGVTCQLTQKAMPEGVKLPSTYKALNSNTQVSGRKRREPEVAEALAQDN--VI 892

QY 1006 NCSIAVCORIQCDDIPFFGIOEENAT-----LKGNLSFDWYIKTSHNHLIIVSTAEILF 1059
Db 893 YCASDSCVLINCTI-----DEINASKSVVRLGRF---W-----ERTF 928

QY 1060 NDSVFTLLPQGGAFVRSQETKVE--PFEVNP-----LP-----1092
Db 929 QKAVSELTPVQVQATIASASAANKTIPYNIPLPRDFSDSTKASTLVTBELVPPVPIAW 988
```

QY 1093 -LIVGSSVGGLLLLALITAAKYKIGFFKR-----QYKDMMS--EGGPP 1132
Db 989 WIIWVSVLGGIILLIILILGLWACGFFERKKPCGEKEKYPAPVASADKGGPP 1038

Search completed: November 9, 2004, 12:19:40
Job time: 33.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds
(without alignments)
4783.919 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_02.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852.5	99.6	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4476	76.2	1153	1 ITAM_MOUSE	P05555 mus musculus
3	4387	74.7	1151	2 Q9J130	Q9J130 rattus norv
4	3815.5	64.9	1920	2 Q28984	Q28984 sus scrofa
5	3483	59.3	1163	1 ITAM_HUMAN	P20702 homo sapien
6	3417	58.2	1162	1 ITAD_HUMAN	Q13349 homo sapien
7	3335.5	56.8	1188	2 Q6KAS4	Q6KAS4 mus musculus
8	3335.5	56.8	1188	2 BAD21383	Bad21383 mus muscu
9	3326.5	56.6	1169	1 ITAX_MOUSE	Q9qxn4 mus musculus
10	3251.5	55.3	1161	1 ITAD_RAT	Q9gye7 rattus norv
11	1578	26.9	1165	1 ITAL_BOVIN	P61625 bos taurus
12	1552.5	26.4	1170	1 ITAL_HUMAN	P20701 homo sapien
13	1552	26.4	1166	2 Q6TVB8	Q6TVB8 bos taurus
14	1552	26.4	1166	2 AAQ90015	Aaq90015 bos tauru
15	1533.5	26.1	1163	1 ITAL_MOUSE	P24063 mus musculus
16	1529.5	26.0	1161	2 Q9WTV4	Q9WTV4 mus musculus
17	1519	25.9	1160	2 Q9R200	Q9R200 mus musculus
18	1414	24.1	1196	2 Q9RTF1	Q9RTF1 cyprinus ca
19	1363.5	23.2	1187	2 Q9RTF0	Q98tf0 cyprinus ca
20	1350.5	23.0	1086	2 Q96HB1	Q96hb1 homo sapien
21	1270	21.6	927	2 Q8HZV0	Q8hzv0 bos taurus
22	1163.5	19.8	1167	2 Q8G340	Q8g340 rattus norv
23	1157.5	19.7	1167	1 ITAE_MOUSE	Q60677 mus musculus
24	1149	19.6	1179	1 ITAE_HUMAN	P38570 homo sapien
25	1119	19.0	1167	2 Q88341	Q88341 rattus norv
26	1102.5	18.8	1151	1 ITA1_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1 ITAH_HUMAN	Q9uux5 homo sapien
28	1084	18.5	1170	1 ITAZ_BOVIN	P53710 bos taurus
29	1072	18.2	1178	1 ITA2_MOUSE	Q62469 mus musculus
30	1072	18.2	1178	2 Q6P1C7	Q6pic7 mus musculus
31	1072	18.2	1178	2 AAH5139	Aah5139 mus muscu

ALIGNMENTS

RESULT 1

ITAM_HUMAN	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
ID	P11215;			
AC	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor)			
GN	Names=ITGAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

P17301 homo sapien
O75578 homo sapien
P18614 rattus norv
P61622 mus musculus
Q7tcg3 mus musculus
O42094 gallus galli
Q8bs01 mus musculus
Q9wuf8 mus sp. itg
Q8mkf4 felis silve
Q8tes5 homo sapien
Q9bpq8 halocynthia
Q63001 rattus norv
Q6p966 mus musculus
AAH57200 mus muscu

32 1069 18.2 1181 1 ITA2_HUMAN
33 1065.5 18.1 1167 1 ITAG_HUMAN
34 1060 18.0 1180 1 ITA1_RAT
35 1052.5 17.9 1188 1 ITAH_MOUSE
36 1052.5 17.9 1188 2 Q7TCQ3
37 1051 17.9 1171 2 O42094
38 1041 17.7 1038 2 Q8BS01
39 1041 17.3 895 2 Q9WUF8
40 1013.5 17.3 1160 2 Q8MKF4
41 865 14.7 348 2 Q8TES5
42 852 14.5 1332 2 Q9BPQ8
43 807 13.7 205 2 Q63001
44 784.5 13.4 304 2 Q6P966
45 784.5 13.4 304 2 AAH57200

RT complement receptor type 3 and cellular adhesion molecule Mac-1." ;
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).
 RN [4]
 RP SEQUENCE OF 17-28.
 RX MEDLINE=85188276; PubMed=3887182;
 RA Springer T.A., Teplov D.B., Dreyer W.J.;
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion
 glycoproteins and unexpected relation to leukocyte interferon." ;
 RL Nature 314:540-542 (1985).
 CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various
 adhesive interactions of monocytes, macrophages and granulocytes
 as well as in mediating the uptake of complement-coated particles.
 CC It is identical with CR-3, the receptor for the iC3b fragment of
 CC the third complement component. It probably recognizes the R-G-D
 CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for
 CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides
 CC of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in
 CC mast cell development and in immune complex-mediated
 CC glomerulonephritis. Mice expressing a null mutation of the alpha-M
 CC subunit gene demonstrate increase in neutrophil accumulation, in
 CC response to a impaired degranulation and phagocytosis, events that
 CC apparently accelerate apoptosis in neutrophils. These mice develop
 CC obesity.
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M
 CC associates with beta-2.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=P05555-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P05555-2; Sequence=VSP_010473;
 CC Note=No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
 CC granulocytes.
 CC -!- DOMAIN: The integrin I-domain (Inser) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07640; CAA30479.1; -;
 DR EMBL; AK039444; BAC30350.1; -;
 DR EMBL; M14293; AAA39484.1; -;
 DR PIR; S00551; S00551.
 DR HSSP; P11215; 1BHQ.
 DR MGP; MGI:96607; Itgam.
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.
 DR GO; GO:0007155; P:cell adhesion; IMP.
 DR GO; GO:0045123; P:cellular extravasation; IMP.
 DR GO; GO:0030593; P:neutrophil chemotaxis; IMP.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3-
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWA; 1.
 DR PRINTS; PR01185; INTEGRINA.
 DR PRINTS; PR00453; VWFADOMAIN.
 DR SMART; SM00191; Int_alpha; 5.
 DR SMART; SM00327; VWA; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS00234; VWFA; 1.
 KW Alternative splicing; Calcium; Cell adhesion;
 KW Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;
 KW Signal; Transmembrane.

FT SIGNAL	1	16	
FT CHAIN	17	1153	Integrin alpha-M.
FT DOMAIN	17	1105	Extracellular (Potential).
FT TRANSMEM	1106	1129	Potential.
FT DOMAIN	1130	1153	Cytoplasmic (Potential).
FT REPEAT	31	84	FG-GAP 1.
FT REPEAT	85	163	FG-GAP 2.
FT DOMAIN	164	350	VWFA.
FT REPEAT	337	400	FG-GAP 3.
FT REPEAT	401	452	FG-GAP 4.
FT REPEAT	454	515	FG-GAP 5.
FT REPEAT	517	575	FG-GAP 6.
FT REPEAT	580	632	FG-GAP 7.
FT CA_BIND	465	473	Potential.
FT CA_BIND	529	537	Potential.
FT CA_BIND	592	600	Potential.
FT SITE	1132	1136	GFFKR motif.
FT DISULFID	66	73	By similarity.
FT DISULFID	105	123	By similarity.
FT DISULFID	654	711	By similarity.
FT DISULFID	770	776	By similarity.
FT DISULFID	999	1023	By similarity.
FT DISULFID	1028	1033	By similarity.
FT CARBOHYD	58	58	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	86	86	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	391	391	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	696	696	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	734	734	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	772	772	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	801	801	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	881	881	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	907	907	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	941	941	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	980	980	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	994	994	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	1022	1022	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	1045	1045	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	1051	1051	N-linked (GlcNac. .) (Potential).
FT CARBOHYD	1076	1076	N-linked (GlcNac. .) (Potential).
FT VARSPLIC	453	569	Missing (in isoform 2). /FTId=VSP_010473.
FT CONFLICT	37	37	N -> S (in Ref. 2).
FT CONFLICT	683	683	V -> G (in Ref. 2).
FT CONFLICT	1153	AA; 127480	MW; 178DB988AECB0343 CRC64;
FT SEQUENCE	1153	AA; 127480	MW; 178DB988AECB0343 CRC64;

Query Match 76.28; Score 4476; DB 1; Length 1153;
 Best Local Similarity 74.37; Pred. No. 2.4e-287;
 Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCBPI	60
Db	17	FNLDTEHPMTFQENAKGFGQNVVLGTSVVVAAPQAKAVNQTGALYQCDYSTSRCHPI	76
Qy	61	RLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHQTSENTRYKGLCFELFGSNLRQDPQK	120
Db	77	PLQVPPEAVNMSLGLSLAVSTVPQQLLACGPTVHQNCKENTYVNGLCVLFGLSNLRPQQ	136
Qy	121	FPEALRCPOEDSDIAPLIDSGSIIIPDPERMKEFVSTWMEQLKSKTLPFLMOYSEEF	180
Db	137	FPEALRECPQOESDVELIDPGSGSINNIDFQMKKEFVSTWMEQPKSKTLPFLMOYSEEF	196
Qy	181	RIHFTFEFQNNPRLSVKPIPTOLLGRTHATCVRVIRELLNITNGARKNAFKILIVI	240
Db	197	RIHFTFNDFKNSPSRSHVSPIKQLNGRTKTASIRKVVVRELPHKTNGARENAKILVVI	256
Qy	241	TGGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKRGQELUNTATSKPRDHVFOVN	300
Db	257	TGGEKFGDPLDYKDVIEADREGVIRVIVGVGNAPKPSQRRELDTIASKPAGHVFQVD	316
Qy	301	NFEALKTIQNLREKIFAIETGOTGSSSFEHEMSQEGFSAATISNGPLLTSTVGSYDWAG	360
Db	317	NFEALNTIQNLQOEKIFAIETGOTGSSSFEHEMSQEGFSAITISNGPLLTSTVGSYFDWAG	376

CC RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "CDNA cloning and complete primary structure of the alpha subunit of a
RT leukocyte adhesion glycoprotein, p150,95.";
RL ENBO J. 6:4023-4028(1987).
CC [2]
CC RX SEQUENCE FROM N.A.
CC RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95
RT molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
CC [3]
CC RX ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RL J. Biol. Chem. 265:12750-12751(1990).
CC [4]
CC RX SEQUENCE FROM N.A.
CC RX TISSUE=Blood;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailis D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC [5]
CC RX SEQUENCE OF 20-43.
CC RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
CC interaction during inflammatory responses. It is especially
CC important in monocyte adhesion and chemotaxis.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and
CC granulocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81695; AAA59180.1; -.
DR DR EMBL; M29165; -; NOT ANNOTATED CDS.
DR EMBL; M29487; AAA51620.1; ALT SEQ.
DR EMBL; M29482; AAA51620.1; JOINED.
DR EMBL; M29483; AAA51620.1; JOINED.
DR EMBL; M29484; AAA51620.1; JOINED.
DR EMBL; M29485; AAA51620.1; JOINED.
DR EMBL; M29486; AAA51620.1; JOINED.
DR EMBL; BC038237; AAH38237.1; -.
DR PIR; A35584; RWHUIC.
DR GDB; IN3Y; X-ray; A=141-338.
DR Genew; HGNC:6152; ITGAX.
DR MIM; 151510; -.
DR GO; GO:0008305; C:integrin complex; TAS.
DR GO; GO:0004872; F:receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0009887; P:organogenesis; TAS.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWFA; 1.
DR PRINTS; PRO1185; INTEGRINA.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;
KW Glycoprotein; Integrin; Integrin; Magnesium; Polymorphism; Receptor; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 19 Integrin alpha-X.
FT CHAIN 20 1163 Extracellular (Potential).
FT DOMAIN 20 1107 Potential.
FT TRANSMEM 1108 1128 Cytoplasmic (Potential).
FT DOMAIN 1129 1163 FG-GAP 1.
FT REPEAT 34 87 FG-GAP 2.
FT REPEAT 88 ?
FT DOMAIN 165 351 VWFA.
FT REPEAT 402 401 FG-GAP 3.
FT REPEAT 402 453 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 518 576 FG-GAP 6.
FT REPEAT 581 633 FG-GAP 7.
FT CA_BIND 466 474 Potential.
FT CA_BIND 530 538 Potential.
FT CA_BIND 593 601 Potential.
FT SITE 1131 1135 GFPR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 655 712 By similarity.
FT DISULFID 771 777 By similarity.
FT DISULFID 848 863 By similarity.
FT DISULFID 998 1022 By similarity.
FT DISULFID 1027 1032 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).
FT VARIANT 48 48 W->R (in dbSNP:11574633).
FT FTID=VAR 018672.
FT CONFLICT 209 209 T->S (in Ref. 4).
FT CONFLICT 251 251 T->A (in Ref. 4).
FT CONFLICT 469 469 T->S (in Ref. 4).
FT CONFLICT 490 490 G->A (in Ref. 2).
FT CONFLICT 547 547 E->K (in Ref. 4).
FT CONFLICT 756 756 D->L (in Ref. 1).
FT CONFLICT 819 819 I->V (in Ref. 4).
FT CONFLICT 1161 1163 SEK -> TPHYQDNV (in Ref. 4).
FT STRAND 150 157

```
FT TURN 160 161
FT HELIX 164 178
FT TURN 179 180
FT TURN 183 185
FT STRAND 186 193
FT STRAND 197 201
FT HELIX 203 208
FT HELIX 212 216
FT TURN 217 218
FT STRAND 226 226
FT HELIX 228 236
FT TURN 237 240
FT HELIX 242 244
FT TURN 245 245
FT TURN 248 249
FT STRAND 251 258
FT STRAND 263 263
FT HELIX 269 278
FT TURN 279 280
FT STRAND 282 288
FT HELIX 290 293
FT TURN 296 297
FT HELIX 298 304
FT HELIX 310 312
FT STRAND 313 316
FT HELIX 319 325
FT HELIX 326 334
FT TURN 335 335
SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5DE8F CRC64;

Query Match
Best Local Similarity 59.3%; Score 3483; DB 1; Length 1163;
Matches 693; Conservative 138; Mismatches 292; Indels 6; Gaps 4;

Qy 1 PNLDTENAMTPOENARGQGVVQLQGSVVVVGAPQEIIVANRGSIYQCDYSTGSCEPI 60
Db 20 PNLDTTELTAFRVDSAGFGDSVGVYANSVVVVGAPQKITAAQTGGLYQCGYSTGACEPI 79

Qy 61 RLOVPVAVNMVSLGLSLAATTPSQLLACGTVHTQTSNTYVVKGLCFGLFSNLRQPOK 120
Db 80 GLQVPPVAVNMVSLGLSLAATTPSQLLACGTVHTQTSNTYVVKGLCFGLFSNLRQPOK 137

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTKLFLSMQYSEEF 180
Db 138 LPVSRQCEPREQDIFVLIDGSGSISSRNFAWNVFRAVISQRPSTQPSLQFQSNKF 197

Qy 181 RIHTTFEQNNPNRSLVKEITQLGRTHATGVRKVIRELLNITGARKNAFKILIVI 240
Db 198 QTHFTFEFRFTSNPLSLLASVHQLQGFYTAIONVVRHLFHASYGARRDATKILIVI 257

Qy 241 TDGKFGDPLGYEDVPEADREGVIRVYVIGVDAPRSEKSFQELNTTASKPPRDHVFQVN 300
Db 258 TDGKFGDPLGYEDVPEADREGVIRVYVIGVDAPRSEKSFQELNTTASKPPRDHVFQVN 317

Qy 301 NFEALKTIQNLREKIFAIEGTOTGSSSSFFHEMSQGFSAATISNGPLLSITVGSYDWAG 360
Db 318 DFDALKDIONLKEKIFAIEGTETSSSFELEMAQGFSAVFTPDGPVLGAVGSFTWSG 377

Qy 361 GVFLYTSKSKSTFNTMRVDSMDNDAYLGAAAIIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFNMSENQVMDRDSVLGYSTELALMKWQSLVILGAPRYOHTKAVIFT 437

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGGVSCPL 480
Db 438 QVSRQWRKAEVGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGGVSCPL 497

Qy 481 PRGORARQCDVAVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPEGEDNRGAYLYF 540
Db 498 PRGWR-RWVCDVAVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVVIAGPEENRGAYLYF 556

Qy 541 HGTSGSGISPSHSORIASGSKLSPRIQYFGQSLSGQDITMDGLVDLTVGAQCHVILLRSQ 600
Db 557 HGVLGSPISPSHSORIASGSKLSPRIQYFGQSLSGQDITMDGLVDLTVGAQCHVILLRTR 616
```

```
Qy 601 PVLEVKAIMENPREVARNVPECDNDQVYKGEAGEVVRVCLHVQKSTRDRREGOIQSVVT 660
Db 617 PVLWVGVMQFIPAEIIPRSAPFEQREQVVEQTLVQSNICILYIDKRSNLRDLSQSVT 676

Qy 661 YDLALDSGRPHSRAVFNETKSNTRRQTVLGLTCTETLKLQLPNCIEDPVSPVLRINF 720
Db 677 LDALDPGLSPRATFOETKNSLSRVRLGLKAHCENFNLLLPSCVEDSVTPITLRNF 736

Qy 721 SLVGTPLSAGNLAPVLAEDAQRULTALFPFEKNCNDNICODDLSITFFSMSLDCLVVG 780
Db 737 TLVGKPLLAFLNLRPMLAADAQRYFTASLPFEKNCGADHICQDNLGISFSFPGKSLIVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKVKVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 SNLENAEVMVNDGEDSYGTTITFSHPAGLSYRYVAGQKQGRSLHLTCDAPVG-- 854

Qy 841 SGALKSTSCSINHPIFFENSEVENTIITFDVDSKASGLNKLKLLKANVTSENMMPTNKTEF 900
Db 855 SQGTWSTSCRINHLIFRGAQITELATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914

Qy 901 QLELPVKVAVYVTVSHGVSTKYLNFTAS-ENTSRVMQHQYQVSNLQORSIPISLVLVP 959
Db 915 QLELPVKVAVYVTVSSHEQFTKYLNFSEBEKSHVAMHYQVNNLQORDLPVSIWVP 974

Qy 960 VRLNOTVIMDRPOVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDI 1019
Db 975 VELNQEAVMDVESHQPNFSLRCSSEKIAPPASDFLAHQKNPVLOCSITAGLRFRCDV 1034

Qy 1020 PFGIOEFENATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFFVRSQTE 1079
Db 1035 PFSFVQELDFTLLKGNLSPGVRQILQKKVSVVSAEITPDTSVYSQLPQGEAFMRAQTT 1094

Qy 1080 TKVEPFPVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFFKRQYKMMSE 1128
Db 1095 TVLEKYVHNPTPLIVGSSIGGLLLALITAVLYKVGFFKRYKEMMEE 1143

RESULT 6
ITAD HUMAN
ID ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RE MEDLINE=96111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RE MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Spl and Sp3.";
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RE MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
RT beta 2-integrin alpha subunit.";
```

RL Gene 171.291-294 (1996).
 RN [4]
 RP INTERACTION WITH VCAM1.
 RX MEDLINE:99059842; PubMed:9841932;
 RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;
 RT "alphabeta2 integrin is expressed on human eosinophils and functions
 as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-
 1).";
 RT J. Exp. Med. 188:2187-2191 (1998).
 RN [5]
 RP INTERACTION WITH VCAM1.
 RX MEDLINE:99370002; PubMed:10438935;
 RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
 binding interface between I domain and VCAM-1.";
 RL J. Immunol. 163:1984-1990 (1999).
 CC !- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
 CC !- VCAM1. May play a role in the atherosclerotic process such as
 CC clearing lipoproteins from plaques and in phagocytosis of blood-
 CC borne pathogens, particulate matter, and senescent erythrocytes
 CC from the blood.
 CC !- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
 CC associates with beta-2.
 CC !- SUBCELLULAR LOCATION: Type I membrane protein.
 CC !- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell
 CC lines and subsets of peripheral blood leukocytes and strongly on
 CC tissue-specialized cells, including macrophages foam cells within
 CC atherosclerotic plaques, and on splenic red pulp macrophages.
 CC !- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.
 CC !- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U37028; AAB38547.1; -;
 DR EMBL: U40274; AAB60634.1; -;
 DR EMBL: U40275; AAB60635.1; -;
 DR EMBL: U40276; AAB60636.1; -;
 DR EMBL: U40277; AAB60637.1; -;
 DR EMBL: U40278; AAB60638.1; -;
 DR EMBL: U40279; AAB60639.1; -;
 DR EMBL: U40278; AAB60638.1; JOINED.
 DR EMBL: AF187881; AAF62875.1; -;
 DR HSSP: P11215; 1BHQ.
 DR Genbank: HGNC:6146; ITGAD.
 DR MIM: 602453; -;
 DR GO: GO:0008305; C:integrin complex; TAS.
 DR GO: GO:0016337; P:cell-cell adhesion; NAS.
 DR GO: GO:0007160; P:cell-matrix adhesion; NAS.
 DR GO: GO:0006955; P:immune response; NAS.
 DR InterPro: IPR000413; Integrin response; NAS.
 DR InterPro: IPR002035; WVF A.
 DR Pfam: PF01839; FG-GAP; 3
 DR Pfam: PF00357; Integrin_alpha; 1.
 DR Pfam: PF00092; WVA; 1.
 DR PRINTS: PR01185; INTEGRIN.
 DR PRINTS: PR00453; VWFADOMAIN.
 DR SMART: SM00191; Int alpha; 5.
 DR SMART: SM00327; WVA; 1.
 DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE: PS00234; WVF A; 1.
 KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; signal; Transmembrane.
 FT SIGNAL 1 17 Potential.

FT	CHAIN	18	1162	Integrin alpha-D.
FT	DOMAIN	18	1100	Extracellular (Potential).
FT	TRANSMEM	1101	1121	Potential.
FT	DOMAIN	1122	1162	Cytoplasmic (Potential).
FT	REPEAT	32	85	FG-GAP 1.
FT	REPEAT	86	?	FG-GAP 2.
FT	DOMAIN	150	332	VWFA.
FT	REPEAT	350	400	FG-GAP 3.
FT	REPEAT	401	452	FG-GAP 4.
FT	REPEAT	454	516	FG-GAP 5.
FT	REPEAT	518	576	FG-GAP 6.
FT	REPEAT	581	633	FG-GAP 7.
FT	CA_BIND	465	473	Potential.
FT	CA_BIND	530	538	Potential.
FT	CA_BIND	593	601	Potential.
FT	SITE	1127	1131	GFPR motif.
FT	DISULFID	67	74	By similarity.
FT	DISULFID	106	124	By similarity.
FT	DISULFID	655	710	By similarity.
FT	DISULFID	769	775	By similarity.
FT	DISULFID	846	861	By similarity.
FT	DISULFID	994	1018	By similarity.
FT	DISULFID	1023	1028	By similarity.
FT	CARBOHYD	59	59	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	87	87	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	99	99	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	391	391	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	691	691	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	733	733	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	873	873	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	957	957	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1046	1046	N-linked (GlcNAc. .) (Potential).
FT	CONFLICT	500	500	Missing (in Ref. 2).
FT	CONFLICT	515	518	GHFW -> ATP (in Ref. 2).
FT	CONFLICT	825	825	L -> V (in Ref. 2).
FT	CONFLICT	984	984	V -> A (in Ref. 2).
SQ	SEQUENCE	1162 AA;	126885 MW;	F296A1A35455D77D CRC64;

Query Match 58.2%; Score 3417; DB 1; Length 1162;
 Best Local Similarity 59.8%; Pred. No. 3.7e-217;
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
 QY 1 ENLDTENAMTFQSNARGFGQSVVQLQSGRWVVGAPQEIIVAAVORGSLYOCDYTGSCPEI 60
 DB 18 FNLDVEEPTIFQEDAGGFGQSVVQFGSRIVVGAPLEVAANOTGRLYDCAAATGMCOPI 77
 QY 61 RLQVPVEAVNMSIGLSLAATTSPQLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQPOK 120
 DB 78 PLHIRPEAVNMSIGLSLAATTSPQLACGPTVHQTCSNTYVYKGLCFPLFGSNLRQPOK 136
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTKLFSLMQYSEEF 180
 DB 137 VPDATPECPHQEMDVIIVFLIDGSGSIDONDFNMKGFEVQAVMGQFEGTDLTFLALMQYSNLL 196
 QY 181 RHFTPEKQNNPNSRLVKPTQLLGRTHRTATGVKVRKRELINITNGARKNAFKILIVI 240
 DB 197 KHFTFTQFTSFSQSLVDPIVLQKLTFTATGILTVTLFHHKNGARKSAKILIVI 256
 QY 241 TDGEKFGDPLGYEDVTPADREGVIRYIVGVGDAPFSEKSRQELNTIASKPRDHFVQVN 300
 DB 257 TDQKYKDPLEYSVDVIPQAEKAGIIRYAGVGHAFQGPATQELNTISSAPPDHFVKVD 316
 QY 301 NFEALKTIQNLREKIFAECTQTGSSSFEHMSQEGSAITSNGLPSTVGSVDWAG 360
 DB 317 NFAALGSIQKQLEKIYAVEGTQSRASSFQHEMSQEGFSTALTMDOGLFAGVGSFWSG 376
 QY 361 GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAYAILNRNVSQSLVGLAPRYOHLGLVAMER 420
 DB 377 GAFLYPNMSPFTINMSQENVMDRDSYLGSTELALWGVQNLVGLAPRYOHLGLVAMER 436
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQSVQVPL 480
 DB 437 QVSRQWRKKAETGTQIGSVFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQSVQVPL 496

```
QY 481 PRGORARWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540
Db 497 PRGORVQWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 556
QY 541 HGTSGSGISPSHSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLTVGAGHVVLLRSQ 600
Db 557 HGASESGISPSHSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLTVGAGHVVLLRSQ 616
QY 601 PVLRVKAIMENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
Db 617 PVLKVGVMARFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSVR 674
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQCTETKLOLPCNCTEDPVSPTVLINLF 720
Db 675 FDLALDGRPHSRVAFNETKNSRROTQVGLTQCTETKLOLPCNCTEDPVSPTVLINLF 734
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 735 SLVREPIPSQNLRPVLAEDAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 794
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRACBSASTEV 840
Db 795 SSLELNIVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRACBSASTEV 853
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 854 EG-LRSGRCNVHPIFEGSNGTFTVDSYKATLGDRLMIRASASNNKASSKATP 912
QY 901 QLELPVKYAVVMVTSRGVSTKYLNLF-TASENTSRVMQHOVQVSNLQORSIPISLFLVP 959
Db 913 QLELPVKYAVVMVTSRGVSTKYLNLF-TASENTSRVMQHOVQVSNLQORSIPISLFLVP 972
QY 960 VRLNQTVMDRPPQVTFSENLSSTCTYERLPDSHSDFLAELRKAPVNCSTAVCQRIODI 1019
Db 973 VLLNGVAVMDVMEAPSQSL--PCVSRKPKPQHSDFLTQTSRSPMLDCSTADCLQFRCDV 1030
QY 1020 PFGCIQEEFNATLGNLSFDWVIKTSNHNLLIVSTABILEFNDVSFTLLPGQAFVRSOTE 1079
Db 1031 PFSVSQELDFTLGNLSFGWVRETLOKKVLIVSVAEITFDTSVYSQLPQGEAFMRAQME 1090
QY 1080 TKVEPPEVPNPLPLIVGSSVGGLLALLAIIAALYKLGFRKQYKXMMSE 1128
Db 1091 MWLEDEVDVNAIPIIMSSVVGALLLALITATLYKLGFRKHYKMELED 1139
RESULT 7
Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (fragment).
GN Name=MFLJ00114;
OS Mus musculus (mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kituno R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries."
RL DNA Res. 11:167-180(2004).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AK311133; BAD21383.1; --
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; VWFADOMAIN.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS02334; VWA; 1.
DR Cell adhesion; Integrin; Transmembrane.
FT NON TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.8%; Score 3335.5; DB 2; Length 1188;
Best Local Similarity 57.4%; Pred. No. 9.8e-212;
Matches 654; Conservative 165; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARFGQSVVQLOQSGVVGAPQIIVANQSGSLYQCDYSTGSCPEI 60
Db 39 FNLDAREPPTHFMDGAEFGHSVIQYDSSVWVVGAPKEIKATNQIGGLYKCYHTGNCEPI 98
QY 61 RLQVPEAVNMSLGLSLAATTSPQLACGPTVHQTCSNTYKGLCFLEGSNLRQOPQK 120
Db 99 SLQVPEAVNMSLGLSLAATNPSWLLACGPTVHTTCRENIYLTGLCFLLSSFSKQS-QN 157
QY 121 FPEALRCPOEDSDIAPLIDGSGSIIPHDPRMKEFVSTVMEQKKSKTTLFSLMOYSEEP 180
Db 158 FPTAAQCPKQDQDIVFLIDGSGSISSTDFEKLDFVKAVMSQLQRESTRFSLMQFSDYF 217
QY 181 RIHTFKEFQNNPRSLVPIPTOLLGRTHATGVKRVIRELLNITMGARKNAFKILIVI 240
Db 218 RVHTFNNFISTSSPLSLDSVRQLRGYTTASAKHIVITELFTTQSGARQDATKVLIVI 277
QY 241 TDGEKFGDPIGEDVIEPADREGVIRVYIGVDAFRSEKSRQELNTIASKPRDHVQVN 300
Db 278 TDRKQKQDNLSDYSVIEMAEAAAIIRVAIGVKAFFNEHKSQELKALIASMPSEHYVFSVE 337
QY 301 NFPAKTIQNLREKIPAEIGTQGSSESSSEHMSQEGFSAATSGNPLISTVCSYDWAG 360
Db 338 NFDAKDIEQLKEKIPAEIGTQGSSESSSEHMSQEGFSAATSGNPLISTVCSYDWAG 397
QY 361 GVFLYTSKESKTFINTRVDSMDNDAYGAAAIILRNRVQSVLGVAPRYQHIGLVAMFR 420
Db 398 GAFLYPSNMPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTKGVIFT 457
QY 421 QNTGMMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTQTRGGQVSCPL 480
Db 458 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDSSTDLVLIGVPHYEHTRGQVSCPM 517
QY 481 PRGORARWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540
Db 518 P-GVGRRWHCQGTTLHGEGHPWGRFGAALTIVLGDVNGDSLADVAIGAPGEENRGAYVIF 576
QY 541 HGTSGSGISPSHSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLTVGAGHVVLLRSQ 600
Db 577 HGASRQDIAPSPQRISASQIPRIQYFGQSLGGQDLTMDGLVDLVANGSKRVLLRTR 636
QY 601 PVLRVKAIMENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
Db 637 PILRVSPVTHFTPAEITSRSVFCEQVAPETQLSDATVCLHIESPKTQL--GDLRSTVT 694
QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVGLTQCTETKLOLPCNCTEDPVSPTVLINLF 720
Db 695 FDLALDGRPHSRVAFNETKNSRROTQVGLTQCTETKLOLPCNCTEDPVSPTVLINLF 754
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 755 SLVGVPISSLQNLQPMALVDQYFTASLFFKNCAGADHICQDDLSVVFPGPDLKTLVVG 814
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQ-----NQSRQSWR 829
Db 815 SDLELNVDVTVSNDGEDSYGTITVTFYFVGLSFRVAEGQVFLRKKEDQWQRQGHSLH 874
```

```

QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSE 889
DB 875 LMCD--STPDRSQGLWSTSCSRHVI FRGGSQMTFLVTFDVPKABLGDRLLRLARVGSE 932
QY 890 NNMPTNKTEFOLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 948
DB 933 NNVPCTPKTTFQLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 992
QY 949 SUPISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNC 1008
DB 993 DVPVSINFVPIELKGEAVW-TVMVSHPNQPLTCYRNRLKPTQDFLTHMQKSPVLDGS 1051
QY 1009 IAVCQRIQCDIPFFGIQEEFNATLKNLSDFWIKTSHNHLIVSTAELFNDVSFTLLP 1068
DB 1052 IADCLHLRCDIPSLGILDLYFLKGNLSFGWISQTLQKKVLLSEAEITFNTSVYQLP 1111
QY 1069 GQAFVRSQTEKVEPFEVNPPLVIVGSSVGGLLALITAAIYKLGFFKQYKDMMS 1128
DB 1112 GQAEFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLALITAAIYKLGFFKQYKEMLEE 1171

RESULT 8
BAD21383
AC BAD21383 PRELIMINARY; PRT; 1188 AA.
DT 01-JUN-2004 (T-EMBLrel. 27, Created)
DT 01-JUN-2004 (T-EMBLrel. 27, Last sequence update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, AKL31133; BAD21383.1; -.
FT NON_TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match
Best Local Similarity 57.4%; Pred. No. 9.8e-212; Length 1188;
Matches 654; Conservative 165; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAAVQVSLYQCDYSTGSCPEI 60
DB 39 FNLDAEKPTFHFDGABFGHVSQVQDSSWVVGAPKEIKATNQIGGLYKCGYHTGNCPEI 98
QY 61 RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHTQTCSENTYVKGICFLFGSNLRQOPQK 120
DB 99 SLQVPPEAVNMSLGLSLAATNSPWLACGPTVHTTCRENIYLTGLCFLSSSFQKS-QN 157
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKFVSTVMEQLKSKSLTFLSLMOYSEEF 180
DB 158 FPTAQECQKQDQIVFLIDGSGISSTDFEKMLDFVKAVMSQLQRPSTFSLMQSDYF 217
QY 181 RIHFTKFEQNNPRSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
DB 218 RVHFTFNFTSTSPSLSDSVQLRGYTYTASAIKHVITELFTTQSGARQDQATKVLVI 277
QY 241 TDGEKGDPLGYEDVPEADREGVIRYVIGVGAFAFSEKSRQELINTIASPPRDHVFQVN 300
DB 278 TDGRKQGDNLSDYSDVIPMAEASIIIRYVIGVGAFAFNEHSEKQELKATAMPSHEYVFSVE 337
QY 301 NFEALTKIQNLREKIFAIGTQTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360

```

RESULT 9

```

ID ITAX MOUSE STANDARD; PRT; 1169 AA.
AC Q9QX4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```

```

DB 338 NFDALKDIOENQKLEKIFALEGTETSPSSSTFELEMSQEGESAVFTPDGVLGAUGSFSW 397
QY 361 GVFLYTSKESKTFINNRVDSMDNDAYLGYAAAILRNVRQSVLVLGAPRYQHIGLVAMER 420
DB 398 GAFLYPSNMRPTFINMSQENEDMDAYLGYSTALAFWKGVHSLILGAPRHQHTKGVVIFT 457
QY 421 QNTGWESNANVKGTOIGAYFGASLCSDVDSDNGSTDLVLI GAPHYVEOTRGGOVSCPL 480
DB 458 QSRHRWPKSEVRGTOIGSYFGASLCSDVMDRGGSTDLVLI GVPHYEYTRGQSVSCPM 517
QY 481 PRGQARWOCDAVLGYEQCPWGRFGAALTVLGDVNGDKLTDVAI GAPGEEDNRGAVYLF 540
DB 518 P-GVGERWHICGTLHGEQHPWGRFGAALTVLGDVNGDSLADVAI GAPGEENRGAVYIF 576
QY 541 HGTSSGSISSHQRIAGSKLSPRIQYFQOSLGGQDLTMDGLVDLTVGAQHVLRLSQ 600
DB 577 HGASRQDIAPSPQRISASQIPRIQYFQOSLGGQDLTRDGLVDLTVGAQSKRVLRLTR 636
QY 601 PVLRYKAIEMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
DB 637 PILRVSPVHTFPAEISRSVFECQEQVAPQTLSDATVCLHIESPKQL--GDLRSTVT 694
QY 661 YDLALDSGRPHSAVFNETKNSTRQTVLGLTOTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 695 FDLALDHGRLSTRAIFKETKTRALTGVKTLGLNKHCSVKLLLPACVEDSVPTILRLNF 754
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALPFFPKNCNDNICODDLITFSFMSLDCLVVG 780
DB 755 SLVGVPISSLQNLQPLAVDDQTYFTASLPFFKNCAGADHICODDLISWVFGFPLDKTLVVG 814
QY 781 GPREENVTYVRNDCGDSYRTQVTFEPLDLSYRKVSTLQ-----NORSORSWR 829
DB 815 SDELENVDTVSDGSDSYTGVTFVPLVGLSFRVABEQVFLRKEDQOMORRQSHLH 874
QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSE 889
DB 875 LMCD--STPDRSQGLWSTSCSRHVI FRGGSQMTFLVTFDVPKABLGDRLLRLARVGSE 932
QY 890 NNMPTNKTEFOLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 948
DB 933 NNVPCTPKTTFQLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 992
QY 949 SUPISLVFLVPVRLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNC 1008
DB 993 DVPVSINFVPIELKGEAVW-TVMVSHPNQPLTCYRNRLKPTQDFLTHMQKSPVLDGS 1051
QY 1009 IAVCQRIQCDIPFFGIQEEFNATLKNLSDFWIKTSHNHLIVSTAELFNDVSFTLLP 1068
DB 1052 IADCLHLRCDIPSLGILDLYFLKGNLSFGWISQTLQKKVLLSEAEITFNTSVYQLP 1111
QY 1069 GQAFVRSQTEKVEPFEVNPPLVIVGSSVGGLLALITAAIYKLGFFKQYKDMMS 1128
DB 1112 GQAEFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLALITAAIYKLGFFKQYKEMLEE 1171

```

RA Huang X., Gorski K., Tong C., Rattis F.-M., Teeng S.-Y., Pardoll D.,
 RA Tsuchiya H.;
 RL "Isolation of genes selectively expressed by dendritic cells.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell
 CC interaction during inflammatory responses. It is especially
 CC important in monocyte adhesion and chemotaxis (By similarity).
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X
 CC associates with beta-2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
 CC with I-domains do not undergo protease cleavage.
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -!- SIMILARITY: Contains 1 VWFA domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF211864; AAP23492.1; -;
 DR HSSP; P20702; IN3Y.
 DR MGI; MGI:96609; Itgax.
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00357; Integrin_alpha; 1.
 DR Pfam; PF00092; VWF; 1.
 DR PRINTS; PRO1185; INTEGRIN.
 DR PRINTS; PRO0453; VWFADOMAIN.
 DR SMART; SM00191; Int alpha; 5.
 DR SMART; SM00327; VWF; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 DR PROSITE; PS50234; VWFA; 1.
 DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 19 Potential.
 FT CHAIN 20 1169 Integrin alpha-X.
 FT DOMAIN 20 1116 Extracellular (Potential).
 FT TRANSMEM 1117 1137 Potential.
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).
 FT REPEAT 34 87 FG-GAP 1.
 FT REPEAT 88 ? FG-GAP 2.
 FT DOMAIN 152 330 VWFA.
 FT REPEAT ? 402 FG-GAP 3.
 FT REPEAT 403 454 FG-GAP 4.
 FT REPEAT 456 518 FG-GAP 5.
 FT REPEAT 519 577 FG-GAP 6.
 FT REPEAT 582 634 FG-GAP 7.
 FT CA BIND 467 475 Potential.
 FT CA BIND 531 539 Potential.
 FT CA BIND 594 602 Potential.
 FT SITE 1140 1144 GPFKR motif.
 FT DISULFID 69 76 By similarity.
 FT DISULFID 108 126 By similarity.
 FT DISULFID 656 711 By similarity.
 FT DISULFID 770 776 By similarity.
 FT DISULFID 858 873 By similarity.
 FT DISULFID 1007 1031 By similarity.
 FT DISULFID 1036 1041 By similarity.
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 267 267 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 949 949 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1059 1059 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 1084 1084 N-linked (GlcNAc...) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033219A6 CRC64;
 Query Match 56.6%; Score 3326.5; DB 1; Length 1169;
 Best Local Similarity 57.2%; Pred. No. 3.8e-211;
 Matches 652; Conservative 167; Mismatches 302; Indels 19; Gaps 7;
 Qy 1 FNLDTNAMTFQENARGCGSVVQLQSGRVVVGAPQEIIVANORGSLVQCDYSTGSCPEI 60
 Db 20 FNLDAEKLTHFMDGAEFGHSLVLYQDSVVVVGAPKEIKATNQIGGKYKGYHTGNCPEI 79
 Qy 61 RLQVPEAVNMSLGLSLAATSPPLQACGPTVHTQCTSENTYVKGCLFGLFSGNLRQOPQK 120
 Db 80 SLOVPPRAVNISLGLSLAATNPSPWLLACGPTVHTCRENIYLTGLCELLSSPKQS-QN 138
 Qy 121 FPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQKKSKTKTLFSLMOYSEEF 180
 Db 139 FPTAQACPKQDQDIVELIDGSGSISSTDFEKMDFVKAVMSQLQRPSTRESLMQFSDYF 198
 Qy 181 RIHFTFEFONNPNRSLVPIPTOLLGKTHATGVKRVIRELLNITGARKNAFKILIVI 240
 Db 199 RVHFTFNNFISTSPSLSGSVRLRGYTYTASAKHIVITELFTTQSGARQDATKVLIVI 258
 Qy 241 TDCEKFGDPLGYEDVIEADREGVIRYVIGVDARSEKSRQELNTIASKPPRDHVFQVN 300
 Db 259 TDCRKQGDNLSDYSVIMPAEAAAIIRVAIGVKAFYNEHSKQELKAIASMPSEHVSFVE 318
 Qy 301 NFEALKTIONLRKIPFAIEGTQTGSSSSFEHEMSQGFSAITSTNPLSTVGSYDWAG 360
 Db 319 NFDALKDIENQLKEKIPFAIEGTETPSSSTFELEMSQGFSAVFTPDGVLGAVGSFWSG 378
 Qy 361 GVFLYTSKEKSTFINMTFRVSDMNDAYLVGAATAIILNRVQSLVILGAPRYOHIGLVAMER 420
 Db 379 GAFLYPSNMRPTFINMSQENEDMRDALVLTALAFWKGVHSLILGAPRHQHTOKVIFT 438
 Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSGSTDVLVIGAPHYEYQTRGGQVSCPL 480
 Db 439 QESRHRPKGEVRGTQIGSYFGASLCSDVMDRDGSTDVLVIGVPHYEYTRGGQVSCPM 498
 Qy 481 PRGQARWQCDVLYGEGOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLIF 540
 Db 499 P-GVGRWKGHTTLLHGEQHPGRFGAALTVLGDVNGDSLADVAIGAIGAPGEENRGAYIIF 557
 Qy 541 HGTSGSGISPSHSORAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAGQHVLILRSQ 600
 Db 558 HGASRQDIAPSPQRIASASQIPRIQYFGQSLGGQDLTRDGLVDLAVGSKGRVLLLR 617
 Qy 601 FVLRVKAIMEFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGIQISVVT 660
 Db 618 PILRVSPVHTPAEISRSVFECQEQVAPETLLSDATVCLHIHESPKTQL--GDLRSTVT 675
 Qy 661 YDLALDSGRPHSRAVENETKNSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILNLF 720
 Db 676 FDIALDHGRSLSTRAIFKETKTRALT RVKTLGLKNKCESVKLLPACVEDSVPTILNLF 735
 Qy 721 SLVGTPLSAFENLRPVLAEADAQRALTALFPFKKNCNDNI CQDDLSTFTFSMDCLVVG 780
 Db 736 SLVGVPISSLQNLQPLAVDDQYTFIASLFPKNCAGADHICQDDL SVVFGFDPDLKTLVVG 795
 Qy 781 GPREFNVTVTRNDGEDSYFTQVTFPPDLDSYRKVSTLQ-----NORSQSRWR 829
 Db 796 SDLELNVDVTVNSDGEDSYGTVTTLFPYVGLSPRRVAEGQVFLRKKEKEDQWQRGQSLH 855
 Qy 830 LACESASSTVSGALKSTSCSINHPTFPENSEVTENITFDVDSKASIGNKLLKANVTSE 889
 Db 856 LMCD--STPRSOGLWSTSCSRHVIFRGSGQMTFLVTFDVSRAELGDLLRLRARVGE 913
 Qy 890 NNMPRTNKTEFQLELPVKYAVVMVTSKYLNFASE--NTRSVMQHQYQVSNLQGR 948
 Db 914 NNVPGPKTTFQLELPVKYAVYTWISSHDQFTKYNLNFSTSEKETSVEHFRFQVNNLQGR 973
 Qy 949 SLPISLVFLVPRLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNCVS 1008
 Db 974 DVPVSINFVPIELKGEAVW--TVMVSHPNPLTQCYRNRLKPTQFDLLTHMOKSPVLDCS 1032

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y00796; CAA68747.1; --
 EMBL; AC002310; AAC31672.1; --
 PIR; S03308; S03308.
 PDB; 1CQP; X-ray; A/B=153-334.
 PDB; 1DQG; NMR; A=149-336.
 PDB; 1LFA; X-ray; A/B=--
 PDB; 1MJN; X-ray; A=153-331.
 PDB; 1MQ8; X-ray; B/D=155-331.
 PDB; 1MQ9; X-ray; A=152-330.
 PDB; 1MQA; X-ray; A=152-330.
 PDB; 1ZON; X-ray; A=150-336.
 PDB; 1ZOO; X-ray; A/B=150-336.
 PDB; 1ZOP; X-ray; A/B=150-336.
 Genew; HGNC:6148; ITGAL.
 MIM; 153370; --
 GO; GO:000805; C:integrin complex; TAS.
 GO; GO:0006928; P:cell motility; TAS.
 InterPro; IPR000413; Integrin_alpha.
 InterPro; IPR002035; VWF A.
 Pfam; PF01839; FG-GAP; 3.
 Pfam; PF00357; Integrin_alpha; 1.
 Pfam; PF00092; VWA; 1.
 PRINTS; PR01185; INTEGRINA.
 PRINTS; PR00453; VWFADOMAIN.
 SMART; SM00191; Int_alpha; 5.
 SMART; SM00327; VWA; 1.
 PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 PROSITE; PS0234; VWFA; 1.
 3D-structure; Alternative splicing; Calcium; Cell adhesion;
 Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
 Receptor; Repeat; Signal; Transmembrane.
 SIGNAL 1 25
 CHAIN 26 1170 Integrin alpha-L.
 DOMAIN 26 1090 Extracellular (Potential).
 TRANSEM 1091 1111 Potential.
 DOMAIN 1112 1170 Cytoplasmic (Potential).
 REPEAT 42 91 FG-GAP 1.
 REPEAT 92 149 FG-GAP 2.
 DOMAIN 156 327 VWFA.
 REPEAT 350 400 FG-GAP 3.
 REPEAT 401 455 FG-GAP 4.
 REPEAT 457 516 FG-GAP 5.
 REPEAT 518 575 FG-GAP 6.
 REPEAT 578 630 FG-GAP 7.
 CA_BIND 468 476 Potential.
 CA_BIND 530 538 Potential.
 CA_BIND 590 598 Potential.
 SITE 1115 1119 GFFKR motif.
 DISULFID 73 80 By similarity.
 DISULFID 111 129 By similarity.
 DISULFID 653 707 By similarity.
 DISULFID 771 777 By similarity.
 DISULFID 845 861 By similarity.
 DISULFID 998 1013 By similarity.
 DISULFID 1021 1052 By similarity.
 CARBOHYD 65 65 N-linked (GlcNAc...) (Potential).
 CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
 CARBOHYD 649 649 N-linked (GlcNAc...) (Potential).
 CARBOHYD 670 670 N-linked (GlcNAc...) (Potential).
 CARBOHYD 726 726 N-linked (GlcNAc...) (Potential).
 CARBOHYD 730 730 N-linked (GlcNAc...) (Potential).
 CARBOHYD 862 862 N-linked (GlcNAc...) (Potential).
 CARBOHYD 885 885 N-linked (GlcNAc...) (Potential).
 CARBOHYD 897 897 N-linked (GlcNAc...) (Potential).
 CARBOHYD 1060 1060 N-linked (GlcNAc...) (Potential).
 CARBOHYD 1071 1071 N-linked (GlcNAc...) (Potential).

Q -> QGVHGLVEMOTSKOILCRPAGDAEHTVGAQEGELPC
 PMGVSEAFRDNIRAGPCR (in isoform 2).
 R -> W (in Ref. 1 and 2).
 Y -> I (in Ref. 2).

FT VARSPLIC 954 954
 FT 214 214
 FT CONFLICT 660 660
 FT STRAND 155 162
 FT STRAND 164 164
 FT TURN 165 166
 FT HELIX 169 185
 FT TURN 186 188
 FT STRAND 191 198
 FT STRAND 202 206
 FT HELIX 208 214
 FT HELIX 217 221
 FT TURN 222 223
 FT STRAND 229 229
 FT HELIX 233 243
 FT TURN 244 244
 FT HELIX 247 249
 FT TURN 250 250
 FT TURN 253 254
 FT STRAND 256 263
 FT HELIX 274 276
 FT TURN 277 278
 FT STRAND 280 286
 FT HELIX 288 290
 FT HELIX 293 297
 FT TURN 298 299
 FT HELIX 300 302
 FT HELIX 307 310
 FT STRAND 311 314
 FT HELIX 317 319
 FT HELIX 323 330
 SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.4%; Score 1552.5; DB 1; Length 1170;
 Best Local Similarity 34.5%; Pred. No. 1.4e-93;
 Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;

Qy 1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRRVVVGAPQEIIVANQORSLYQCDYSTGSCE 58
 Db 26 YNLDRVGARFSFPFRAGRHFGRVLRVQV--GNGVIVGAPGE---GNSGTSLYQCQSGTGCHL 81
 Qy 59 PIRLOVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTVVKGLCFELGNSLR--- 115
 Db 82 PVLIR--GSNTSKYLGMILATDPTDGSILACDGLSRTCDQNTYLSGLCYLFRNLQGP 140
 Qy 116 -QQQKPEALRGCPQSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQKLKSKTLFSLM 174
 Db 141 LQGRPGFQECIKG---NVDLVFLFDGSMISLPDPEFQKILDFFMKDVMKLSNTSYQFAAV 196
 Qy 175 QYSEEPRIHFTKFEFQNNPNRSLVKPITQLGRTHATATGVKRVIRELLNITNGARKNAF 234
 Db 197 QFSTSYKTEFDSYVKKRQDPDALLKHVKHMLLTNTFGAINVYVATEVFREELGARPDAT 256
 Qy 235 KILIVITDGEKFGDPLGYEVIPEADREGVIRYVIGVDAPRSKSRKSOELNTIASKPRD 294
 Db 257 KVLIIITDGE--ATDSGNIDAKD-----IIRYIIGIKGFQTKESQETLHKFASKPASE 309
 Qy 295 HVFQVNNFEALKTIQNOLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLPLSTVG 354
 Db 310 FVKILDTFEKLDLFTLEQLKKIVIEGTQSKDLTSFNKELSSSGISADLSRGHAVVGAVG 369
 Qy 355 SYDWAGGVF--LYTSKESKSTINMTRVDSMDNDAYLGAAA--IILNRVQSVILGAPRYQH 412
 Db 370 AKDWAGGFLLDKADLQDDTFIGNEPLTPEVRAGLYGTVTWLPFSRQKTSLLASGAPRYQH 429
 Qy 413 IGLVAMPR--QNTGMESNANVKGTQIGAVFGASLCSDVDVDSNSTDLVLIGAPHYEQT 470
 Db 430 MGRVLLFQEPQGGGHNQSVQTIHGTQLGSYFGGELCGVDVDQDGETELLIGALFLFGEQ 489
 Qy 471 RGGQSVVCPPLPRQGRARWQCDAV--LYGEOQGPWRFGAALTVLGDVNGDKLTDAVIGAP 528

```

Db 490 RGRVFY-----RRQLGFEVSELQDGPYPLGRFGEAITALTIDINGDLVDVAVGAP 544
QY 529 GEDNRGAVYLFHTSGSISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTV 588
Db 545 LEE--QGVYIFNGRG-GLSPQSQRIEQTQVLGSIQWFGRIHGKDLGGDLADVAV 601
QY 589 GAGHVLRLRSQVLKALMEPNPREVARNVFECDQVV-KGKEAGEVRVCLHVOKSTR 647
Db 602 GAEQMVILSSRPVDMVTLMSFSPAIPIVHEVECSYSTSNKMGKGVNITICQI-KSLY 660
QY 648 DLREGIOQSVVYDIALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTELKLQLPNCI 707
Db 661 PQ-OGRLVANLYTYLQDGHRTRRGLPGRGHELRNIAVT-TSMSCDTDFSFHFPVCV 718
QY 708 EDVPSVILRLNFSI---VGTPLS--AFGN-----LRPVLAEDAQRILFTALFFPEKNCGN 757
Db 719 QDLISPINVSLNFSLMEEGTQRDQAQGDIPILRPSLHSETWBI-----BFEKNCGE 773
QY 758 DNICQDLSITTFMSLDCLVGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVS 817
Db 774 DKCEANLVSFSPARSRALRLTAPASLSVELSLNLEEDAYVWQLDHPGGLSRKVE 833
QY 818 TLQNRORSRWLACRS--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875
Db 834 ML---KPHSQIPVSCBELPEESRLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887
QY 876 LGNKLKALKANVTSENN---MPTNKTQFQLELPVKYAVVMVTSHGVS TKYLNFTASEN 931
Db 888 WGSDELHANVTCCNEDSLLENSATTI---IPILYPINILQDQEDSTLYYSFTPKGP 944
QY 932 TSRVMQHOYQV---SNLQORSLP-ISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTK- 986
Db 945 KIHQVKHMVQVRTOPIHDHNIETLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYED 1002
QY 987 -ERLPSHSD--FLAELRKAPVNCVSIACVORIQCIPFFGIGIEFNATLKNLSFDWYIK 1043
Db 1003 LERLPDAAEPCLPALFRCPW-----FRQELVQVIGTLELVEIE 1044
QY 1044 TSHNLLIYSTAILFNDSVFTLLPGGAFVRQTEKVPFVFPNPLPLIVGSSVGGLL 1103
Db 1045 AS-SMFSLSSLSISNSSKHFLYGSNASL-AQVVMKVDDVVEYKQMLYLYLVLSGIGLL 1102
QY 1104 LLALITAAALYKLGFFKROYKDMMESEG-GPPGAP 1136
Db 1103 LLLLLIFVLVKVGFENLKEKMEAGRGVNGIP 1136

RESULT 13
Q6TVB8
ID Q6TVB8 PRELIMINARY; PRT; 1166 AA.
AC Q6TVB8
DT 05-JUL-2004 (TREMblrel. 27, Created)
DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T. Thumrikat P., Kannan M.S., Maheswaran S.K.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AY382558; AAQ9015.1; -.
DR InterPro; IPR004113; Integrin_alpha.
DR InterPro; IPR02035; VWF_A.
DR Pfam; PF01839; FG-GAP; 1.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.

```

```

DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWF; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match      26.4%; Score 1552; DB 2; Length 1166;
Best local Similarity 34.4%; Pred. No. 1.5e-93;
Matches 407; Conservative 200; Mismatches 469; Indels 106; Gaps 33;

QY 1 ENLDTENAMTFOENARG--FGQSVVQLQSGRVVVGAPQEIIVAAQSGSLYQCDYSTGSC 58
Db 25 YNLDVRHVQNFSPPLAGRHFGVRVQV-GNVVVVGAPSE---GNSMGNLYQCOQFETGDC 80
QY 59 PIRLOVPVBAVNMNLSGLSLAATTPPQLLACGPTVHQTCSENTYVRKGLCFGLFSGNL 118
Db 81 PVTLS--SNVTSKYLQMTLATDPTSDNLLACDPLGSRCTCDQNIYLSGLCYLIHENLRGP 138
QY 119 QKPEALRGCPQEDSIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTFLSLMOYSE 178
Db 139 LQHPGQYQICIKGNVDLFLFDGSMLSQDDEFEKIVDFMKDVMKLSNSYSYQFAAQFST 198
QY 179 EPRIHFTFEFQNNPNSLVKPIITOLLGRHTATGVRKVIRELLNITWGARKNAFKILI 238
Db 199 YFTEFTFDLYTKQDPDALLAGVKHMLLTNTFGAINYVAKVFRPDLGARDFATKVL 258
QY 239 VITDGKFGDPLGYEDVPEADREGVIRVYGVGDAPRSEKSEKSELNTIASKPRDRHVQ 298
Db 259 IITDGK--PPTNTTLMRPKTS---RSLIGIKNFKTKESQEAHLHQFASKPVEBFVKI 311
QY 299 VNNFEALKTIQNLREKIFAIEGTQTGSSSSPHEMSQEGSAITNSGNPLLTSTVSYDW 358
Db 312 LDTFEKLKDLFTLEQKIYVIEGTQDLTSFNMELSSSGISADLSEGHGAVGAGAKDW 371
QY 359 AGGVF-LYTSKEKSTFINTRVDSMDNDAVLGYAAA-IILNRVQSLVLCAPRYQHIGLV 416
Db 372 AGFLDLKADLKSSTFVNEPLTVESRAGLYGVTVTRLPSRGTMSLLAGAPKQHVGRV 431
QY 417 AMFRQ--NTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQ 474
Db 432 LLFQPKRGKGPWSQIQEIDGIQIGSYFGELCGVDVDRDGETELLIAAPLYYGEQGRG 491
QY 475 VSVCPPLPRGRARWQCDAVLYGEOGPGRFAGNALTVLGDVNGDKLTDVAIGAPGEDNR 534
Db 492 VFYI---QKIQLEFQMVSELQGETGYPLGRFGAAIAALTIDINGDELTDVAVAPLEE--Q 546
QY 535 GAVYLFHGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHV 594
Db 547 GAVYIFNGQOG-GLSPRPSQRIEQTQVFGSIQWFGRIHGKDLGGDLADVAVAGQGV 605
QY 595 LLLRSQVILRVKAIMFNPREVARNVFECDQVVQKGEAG-EVVRVCLHVQK--STRDLR 651
Db 606 IVLSRPVVDIITVSFSPAETPVHEVECSYSTSNQKEGVNLTVCVQVKSLLIST----F 661
QY 652 EGOIQSVVYDIALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTELKLQLENCIEDPV 711
Db 662 QGHLVANLYTYLQDGHRTSRGLFFGGRKHLIGNTAVTPV-KSCFVFMVFFHFCIQOLI 720
QY 712 SPVLRNLFSL---VGTPLS--AFGNLRPVLAEDAQRILFTALPPFEKNCNDNICODLS 766
Db 721 SPINVSLSLMEEGTTPDRPRALDRDIIPILKPSHLETKEIPFNKCEGDKNCEADLK 780
QY 767 ITFPSMLDCLVVGGRPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKYSTLQNRQSR 826
Db 781 LAFSDMRSKILRLTPSASLSVRLTLTAEDAAYVWQVTLSPFQGLSFRKVEIL--KPHS 837
QY 827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNLKKANV 886
Db 838 HVPVGCEELPEEAHVHS-RALSCNVSPFSGEDSMVDIQVMFNTLQKSGWGDPIELQANV 896
QY 887 TS-----ENNMPRTNKTQFQLELPVKYAVVMVTSHGVS TKYLNFTASENTRVMQH 938

```

Db 897 SCNNEDSLLEDNSGATTS-----IPVMYINVLTKDQENSTLYISFTPKSPRIHHVKH 949
QY 939 QYQV-----SNLQSRSLPISLVLPVRLNQTIV---WD---RPOVTFSS-ENLSSTCHTKE 987
Db 950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLTHKWSIQMEPPVNCSPRNLESFSDAE 1008
QY 988 RLPSSHDFLAELRKAPVNCNSIAVCQRIQCDIPFFGQIEEPNATLKNLSFDWIKTSHN 1047
Db 1009 -----SCSFGT--EPRCPIDF---RQELVQVNGMVELRGITIKAS-S 1044
QY 1048 HLLIVSTAELFNDSVFTLLPGOGAFVRSQTETKVEPFEVNPPLIVGSSVGGLLIAL 1107
Db 1045 MSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLYVEKEMLYLYVLSGIGGLLFL 1103
QY 1108 ITAALYKLGFFKQYKDM-----SEGGPPGABPQ 1137
Db 1104 IFIALLYKVGFFKRLNKEKMEANVDASSEIPGEDAQPELEKE 1145

RESULT 14

AAQ90015
ID AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a."
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY382558; AAQ90015.1; -
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFA896C9DF CRC64;

Query Match 26.4%; Score 1552; DB 2; Length 1166;

Best Local Similarity 34.4%; Pred. No. 1.5e-93;

Matches 407; Conservative 200; Mismatches 469; Indels 106; Gaps 33;

QY 1 FNLDENAMTFQENARG--FCQSVYVLOGSVVVGAPQEIIVAAQNRGSLYQCDYSTGCE 58
Db 25 YNLDRHVQNPFPPLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDL 80
QY 59 PIRLQVPEAVNMISGLSLAATTSPOLLACGPTVHOTCSENTYVKGLCFLFGSNLRQOP 118
Db 81 PVTLSS--SNYTSKYLGMTLATDPTSDNLLACDGLSRTCDQNIYLSGLCYLIHENLRGPV 138
QY 119 QKPEALRGCPQESDIAFLIDGSGSIIPHDFRMKEFVSTVMQMKSKTFLSLMOYSE 178
Db 139 LQHPGQECIKGNVDVFLFDGMSLQQDEFEKIVDFMKDVMMKLSNYSYQFAAVQFST 198
QY 179 EFRTHFTKEPQNNPRSLVKPITQLLGRTHATGVRKVIRELNLITNGARKNAFKILI 238
Db 199 YFRTEFTLDYIKQKDPDALLAGVHKMLLNTGAINYAKVFRPDLGLARPATKVLII 258
QY 239 VITDGEKFGDPLGVEDYEDADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDHFVQ 298
Db 259 IITDGK---PPTNTLMRPKTS---RSLLGIGKNFKTESQEAHLQFASKPVEEFVKI 311
QY 299 VNNFEALTKQNRREKIFALETGQTSSSSFEHMSQEGPSAINTGNPLLSVTGYSDW 358
Db 312 LDTPEKLDKLTQKQIYVTEGTSKQDLTSTFNNMELSSGSIADLSEGHGVVGVAGAKDW 371
QY 359 AGGVF-LYTSKEKSTFTNTMTVDSDMDNDAYLGAAA-IILNRRVQSIVLGAAPRYOHLGLV 416
Db 372 AGGFLDLKADLKSSTFVGNPEPLTVESRAGYLGTVTLRPLSGTWSLLATGAPKYQHVGVR 431

QY 417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSDVDNSNGSTDLVLIGAPHYYQTRGGQ 474
Db 432 LFFQPKRGFPWQIQEIDGIGQISYFGGELCGVDVDRDGETELLTAAPLYYGEQGR 491
QY 475 VSVCLPRGQARQCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPBEENR 534
Db 492 VFYI---QKIQLEFQWSELQGETGYPLGRFGAAIATLDINGDELTDVAVGAPLEE--Q 546
QY 535 GAVYLFHTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGQDGLTMDGLVDLTVGAQOHV 594
Db 547 GAVYIFNGQQG-GLSPRPSQRIEQTQMFQIOWFGRSIHGKDLGGDLGADVAVGAEQV 605
QY 595 LLRSQVLRVKAIMEFNPREVARNPFCNDVVVKKEAG-EVRVCLHVQK--STRDLRL 651
Db 606 IYLSRPVVDLIITSVSFAPIPVHEVECSYSTNQKKEGVNLIVCQVKSLIST----F 661
QY 652 EGQIQSVVYDIALDGRPHSRVAFNETKNSRQTQVLGTQTCETLKQLPNCIEDPV 711
Db 662 QGHLVANLTYTLQDGHRTSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHPFICIQDLI 720
QY 712 SPIVLRNFSL---VGTPLS--AFGNLRPVLAEDAQRLLFTALFFPKNCNDNICODDLS 766
Db 721 SPINVSUSYSLWEEGTDRPRALDRDIPPLIKPSPHLETKEIFPKNCNGDKNCEADLK 780
QY 767 ITFSFMSLDCLVVGGPREFNVTVTRNDGDSYRQTQVTFPPPLDLSYRKVSTLQNRQSR 826
Db 781 LAFSDMSKILRLTPSASLSVLRLNTAEDAYWQVTLSPFQGLSPRKVEIL---KPHS 837
QY 827 SWRIACSASSTEVSGALKSTCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANV 886
Db 838 HVPVGCBLPEEAVVHS-RALSCNVSPFIEDSMVDIQVMFNTLQKSGWGDFTLOANV 896
QY 887 TS-----ENNMPRTNKTQFQLELPVKYAVVMVTVSHGVSTKYLNTASENTRVQMH 938
Db 897 SCNNEDSLLEDNSATTS-----IPVMYINVLTKDQENSTLYISFTPKSPRIHHVKH 949
QY 939 QYQV-----SNLQSRSLPISLVLPVRLNQTIV---WD---RPOVTFSS-ENLSSTCHTKE 987
Db 950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLTHKWSIQMEPPVNCSPRNLESFSDAE 1008
QY 988 RLPSSHDFLAELRKAPVNCNSIAVCQRIQCDIPFFGQIEEPNATLKNLSFDWIKTSHN 1047
Db 1009 -----SCSFGT--EPRCPIDF---RQELVQVNGMVELRGITIKAS-S 1044
QY 1048 HLLIVSTAELFNDSVFTLLPGOGAFVRSQTETKVEPFEVNPPLIVGSSVGGLLIAL 1107
Db 1045 MSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLYVEKEMLYLYVLSGIGGLLFL 1103
QY 1108 ITAALYKLGFFKQYKDM-----SEGGPPGABPQ 1137
Db 1104 IFIALLYKVGFFKRLNKEKMEANVDASSEIPGEDAQPELEKE 1145

RESULT 15

ITAL MOUSE
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=Itgal; Synonyms=Lfa-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1

Db 716 IQDLISPINVSLNLSLEEETPRQKRAMQPIILRPSIHV-TKEIPFEKNCGEDKKCE 774
Qy 763 DDLISITFSFMSLDCLVGGP-----REFNVTIVRNDGEDSYRTQVTFPPPLDLISYRKV 816
Db 775 ANLTLSPPARS-----GPLRLMSSASLAWEWTLNSGEDAYWVRLDLDPFRLSPRKV 827
Qy 817 STLQNRQSRWRILACESASSTEVS GAL-KSTSCSINHPIPPENSEVTFNITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKILKCNVSSPIFKAGQEVSLQWVNTILNSS 882
Qy 876 LGNKLLKANVTSEN-NMPTNKTEFQLELPKYAVVMVVTSHGVSTKYLNFTASENISR 934
Db 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLYPVNILTKEQENSTLIYISFTPKGPKTQ 942
Qy 935 VMQHQYQVSNLQORSPLSILVLPVRLNQTVIWRPQ-----VTFSENLS--TCHTK 986
Db 943 QVOHVYQV-----RIQPSAYDHNMT-LEALVGVPRPHSEDLITYTWSVQTDPLVTCHSE 996
Qy 987 E-RLPSHSDFLAELRKAPVNVNCIAVCORICDIPFGIOEEFNATLKGNLSEFDWYIKTS 1045
Db 997 LKRPSSB---ABQPCLEGV-----QFRCPVIF---RWEILIOVTGTVELSKEIKAS 1042
Qy 1046 HNHLIIVSTABILFNDVSFTLLPGQGAFFVRSQETKVEPFPVNPPLPLIVGSSVGLLLL 1105
Db 1043 -STLSLCSLSVSFNSKHFHYGSKA-SEAQLVAKVDLIHEKEMLVVYVLSGIGGLVLL 1100
Qy 1106 ALITAAALYKLGFFKQYKDM-SEGPPGPAEP 1136
Db 1101 FLIFIALYKVGFFKRNLERMEADGGVPNGSP 1132

Search completed: November 9, 2004, 12:17:43
Job time : 141.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds
(without alignments)
3950.365 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	2	AAW65090 Human Bet
2	5868	99.9	1153	3	AAB07360 Human CD1
3	5868	99.9	1153	5	AAU80252 Human int
4	5868	99.9	1153	5	ABG61469 Human Bet
5	5868	99.9	1153	7	AAO14428 Integrin
6	5868	99.9	1153	7	ADD25615 Binding d
7	5858	99.7	1153	2	AAO41136 Alpha sub
8	5852.5	99.6	1152	8	ADM95959 Human int
9	5845.5	99.5	1152	8	ADP12435 Protein e
10	3487	59.4	1163	8	ADP44061 Human CD1
11	3473	59.1	1163	8	ADNO2004 Human inf
12	3473	59.1	1163	8	ADQ17510 Human sof
13	3464	59.0	1163	2	AAO71120 p150.95 a
14	3450	58.7	1163	2	AAW65091 Human Bet
15	3450	58.7	1163	3	AAB07361 Human CD1
16	3450	58.7	1163	5	ABG61470 Human Bet
17	3448	58.7	1163	6	ABU07406 Protein d
18	3448	58.7	1163	7	ADG32005 Human hom
19	3417	58.2	1161	2	AAU78166 Human bet
20	3417	58.2	1161	2	AAW23049 Human bet
21	3417	58.2	1161	2	AAW57491 Human bet
22	3417	58.2	1161	2	AAW65089 Human Bet
23	3417	58.2	1161	2	AAW72825 Human alp
24	3417	58.2	1161	2	AAW73342 Human alp
25	3417	58.2	1161	3	AAB07359 Human alp

ALIGNMENTS

RESULT 1

AAW65090
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

XX KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
XX KM type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
XX KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;
XX KM rheumatoid arthritis.

XX OS Homo sapiens.

XX PN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX PR 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX DR WPI; 1998-206565/18.

XX PT Screening assay for modulators of integrin binding - using immobilised or
XX PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX PS Example 5; Fig 1A-D; 106pp; English.

XX CC This sequence represents a human beta-integrin CD11b subunit which is
XX CC used to describe a method for identifying compounds that modulate the
XX CC interaction of the beta-integrin alpha-d subunit with a binding partner
XX CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
XX CC -d binding partner, one of which is immobilised and the other of which is
XX CC labelled, in the presence of a test compound, and determining if the
XX CC compound affects binding between the alpha-d polypeptide and alpha-d
XX CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
XX CC comprising the cytoplasmic, transmembrane or extracellular domain of
XX CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

ABG61468 Human Bet
AAW23064 Human bet
AAW65106 Human Bet
Aaw72837 Human alp
Aaw73343 Human alp
Aab07376 Human alp
ABG61485 Human Bet
Aar78169 Rat alpha
AAW23062 Rat beta
AAW60004 Rat alpha
Aaw72824 Rat alpha
Aab07374 Rat alpha
ABG61483 Rat Beta2
AAW65104 Rat beta-
AAW73345 Rat alpha
AAW23061 Mouse bet
AAW60003 Mouse alp
AAW65103 Mouse bet
Aaw72836 Mouse alp
Aaw73347 Mouse alp

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCTSENYYKGLCFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCTSENYYKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRCPCPEDSDIAFLIDGSGSIIIPDFFRMKEFVSTWEOQLKSKTILFSLMOYSEEF 180
DB 137 FPEALRCPCPEDSDIAFLIDGSGSIIIPDFFRMKEFVSTWEOQLKSKTILFSLMOYSEEF 196

QY 181 RIHFTPFQNNPNRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTPFQNNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLIGEDYVIEADREGVIRVIVGVDAPRSEKROELNITIASPPRDHVPQVN 300
DB 257 TDGEKFGDPLIGEDYVIEADREGVIRVIVGVDAPRSEKROELNITIASPPRDHVPQVN 316

QY 301 NFEALKTIQNLREKIFAIETGTGTGSSSPHEHMSQEGFSAAITSGNPLLSITVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIETGTGTGSSSPHEHMSQEGFSAAITSGNPLLSITVGSYDWAG 376

QY 361 GVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQSVCP 496

QY 481 PRGORARWQCDVLYGSGQPGWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF 540
DB 497 PRGORARWQCDVLYGSGQPGWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF 556

QY 541 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSGRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLRVKAIMFENPREVARNVPECNDQVYKGEAGEVRVCLHVQKSTDRLRREGQIQSVVT 660
DB 617 PVLRVKAIMFENPREVARNVPECNDQVYKGEAGEVRVCLHVQKSTDRLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRTOQVLGLTQTCETLKLQLENCHIEDPVSPTVLRINF 720
DB 677 YDLALDSGRPHSRVAFNETKNSRTOQVLGLTQTCETLKLQLENCHIEDPVSPTVLRINF 736

QY 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780
DB 737 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796

QY 781 GPREFNVTIVRNDGDSYQTQVTFPPDLDSYKVKSTLQNSORSWRWLRACASSTEV 840
DB 797 GPREFNVTIVRNDGDSYQTQVTFPPDLDSYKVKSTLQNSORSWRWLRACASSTEV 856

QY 841 SGALKSTSCSINHPIFENSEVETNITFDVDSKASLGNKLLKANVTSENNMPTNKTTEF 900
DB 857 SGALKSTSCSINHPIFENSEVETNITFDVDSKASLGNKLLKANVTSENNMPTNKTTEF 916

QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHVQSVNLGQRS LPISLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMQHVQSVNLGQRS LPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWVNCISIAVCORIQCDIP 1020
DB 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDSFLAELRKAPVWVNCISIAVCORIQCDIP 1036
QY 1021 FFGIQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080
DB 1037 FFGIQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1096
QY 1081 KVPFEPFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
DB 1097 KVPFEPFVNPPLIIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX AAB07360;

DT 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

Human; macrophage infiltration inhibition; alpha_d integrin;
leukocyte integrin; Leu-CAM; leukointegrin; immune response;
inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
lung inflammation; acute respiratory distress syndrome; Crohn's disease;
rheumatoid arthritis; central nervous system injury; CD11b.

OS Homo sapiens.

XX WO200029446-A1.

XX 25-MAY-2000.

XX 16-NOV-1999; 99WO-US027139.

XX 16-NOV-1998; 98US-00193043.

XX 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

XX Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
macrophage infiltration and reduce inflammation at central nervous system
injury sites.

XX Example 5; Fig 1; 270pp; English.

Integrins are a class of membrane-associated molecules that participate
in cellular adhesion. Integrins are made up of an alpha subunit and a
beta subunit. One class of human integrins are restricted to expression
in white blood cells and have a common beta2 subunit: the leukocyte
integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins
have an important role in immune and inflammatory responses. The present
protein sequence is the human integrin alpha subunit CD11b. This sequence
was used in an alignment to identify a novel beta2 integrin alpha
subunit: alpha_d (AAB0014 and AAB07359). The present sequence has
approximately 60% identity to the protein sequence of alpha_d. The
Alpha_d gene and protein may be useful in therapy for diseases linked to
aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple
sclerosis, asthma, psoriasis, lung inflammation, acute respiratory
distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency
(LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the
inhibition of macrophage infiltration at the site of a central nervous
system injury. The monoclonal antibodies can also be used to detect and
diagnose Crohn's disease

SQL	Sequence	1153	AA;
	Query Match	99.9%;	Score 5868; DB 3; Length 1153;
	Best Local Similarity	99.6%;	Pred. No. 0;
	Matches 1133;	Conservative	3; Mismatches 1; Indels 0; Gaps 0;
QY	1	FNLDTENAMTFOENARFGGQSVVOLQSGRVVVGAPQBIIVAANORGSLYQCDYSTGSGCEPI	60
DB	17	FNLDTENAMTFOENARFGGQSVVOLQSGRVVVGAPQBIIVAANORGSLYQCDYSTGSGCEPI	76
QY	61	RLOVPVEAVNMISGLSLAAATSPPLLACGPTVHOTCSENTYVKGCLCFLFGSNLRQOPOK	120
DB	77	RLOVPVEAVNMISGLSLAAATSPPLLACGPTVHOTCSENTYVKGCLCFLFGSNLRQOPOK	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRRMKFEFVSTWMEQLKKSKTILFSLMQYSEEF	180
DB	137	FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRRMKFEFVSTWMEQLKKSKTILFSLMQYSEEF	196
QY	181	RIIHFTPKFQNNPNRSLVPIITQLLQRTTATGVRKVIIRLLNITNGARKNAFKILIVI	240
DB	197	RIIHFTPKFQNNPNRSLVPIITQLLQRTTATGIRKVRLEFNIITNGARKNAFKILIVI	256
QY	241	TDEKFGDPLGYEDVIPADREGVIRVIVGVGDADFSEKSRQELNTTASKPPDRHVFQVN	300
DB	257	TDEKFGDPLGYEDVIPADREGVIRVIVGVGDADFSEKSRQELNTTASKPPDRHVFQVN	316
QY	301	NFEALKTIQNLREKIPAIETGTGGSSSFEHEMSQEGFSAATITSGPLISTVGSYDWAG	360
DB	317	NFEALKTIQNLREKIPAIETGTGGSSSFEHEMSQEGFSAATITSGPLISTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINMTRVDSDMNDAYLGAAAIILNRVQSLVLGAPRYOHTGLVAMFR	420
DB	377	GVFLYTSKEKSTFINMTRVDSDMNDAYLGAAAIILNRVQSLVLGAPRYOHTGLVAMFR	436
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL	480
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQGVSCPL	496
QY	481	PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF	540
DB	497	PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSRPLQYFGQSLGCGQDLTMDGLVDLTVGAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSRPLQYFGQSLGCGQDLTMDGLVDLTVGAQGHVLLRSQ	616
QY	601	PVLVRVKAIEMFNPREVARNYFECNDQVVVKGKEAGEVVCVCHVQKSTRDLREGQIQSVVT	660
DB	617	PVLVRVKAIEMFNPREVARNYFECNDQVVVKGKEAGEVVCVCHVQKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSRVAFNFKTNSTRQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	720
DB	677	YDLALDSGRPHSRVAFNFKTNSTRQVGLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF	736
QY	721	SLVGTPLSAGNLRPVLAEADAQRFTALFPPEKNCNGNDNICQDDLSTTFPSFMSLDCLVWG	780
DB	737	SLVGTPLSAGNLRPVLAEADAQRFTALFPPEKNCNGNDNICQDDLSTTFPSFMSLDCLVWG	796
QY	781	GPREFNVTIVVRNDGEDSYRTQVTFPPDLVSRVKVSTIQNORSQRSWRLACSSASSTEV	840
DB	797	GPREFNVTIVVRNDGEDSYRTQVTFPPDLVSRVKVSTIQNORSQRSWRLACSSASSTEV	856
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTISENNMPTNKTEF	900
DB	857	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTISENNMPTNKTEF	916
QY	901	QLELPVKYAVVMVVTSHGVSTKYKLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV	960
DB	917	QLELPVKYAVVMVVTSHGVSTKYKLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVPV	976
QY	961	RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVMVCSIAVCQRIQCDIP	1020
DB	977	RLNQTVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVVMVCSIAVCQRIQCDIP	1036

[illegible]

SQ	Sequence 1153 AA;	
Query Match	99.9%; Score 5868; DB 5; Length 1153;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 1133; Conservative	3; Mismatches 1; Indels 0; Gaps 0;	
QY	1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 60	
DB	17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI 76	
QY	61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 120	
DB	77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQPOK 136	
QY	121 FPEALRGCPQEDSDIAFLDNGSGIIPDPRRMKEFVSTVMEQLKSKTLFSLMOYSEEF 180	
DB	137 FPEALRGCPQEDSDIAFLDNGSGIIPDPRRMKEFVSTVMEQLKSKTLFSLMOYSEEF 196	
QY	181 RIHFTFEKQNNPNRSLVKPIPTQLLGRTHATGVKVIKRELNITNGARKNAFKILIVI 240	
DB	197 RIHFTFEKQNNPNRSLVKPIPTQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256	
QY	241 TDEKEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRDHVFQVN 300	
DB	257 TDEKEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRDHVFQVN 316	
QY	301 NFEALKTIONLREKIFAEGTQSSSFEHMSQEGFSAITNSGPIILLSTVGSYDWAG 360	
DB	317 NFEALKTIONLREKIFAEGTQSSSFEHMSQEGFSAITNSGPIILLSTVGSYDWAG 376	
QY	361 GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAIILNRNVSQSLVIGAPRYOHIGLVAMPR 420	
DB	377 GVFLYTSKEKSTFINNTRVDSNDMDAYLGAAIILNRNVSQSLVIGAPRYOHIGLVAMPR 436	
QY	421 QNTGMWESNANVKGTQIGAFYFASLCSVDVDSNGSTDVLIGAPHYYEYQTRGQVSVCP 480	
DB	437 QNTGMWESNANVKGTQIGAFYFASLCSVDVDSNGSTDVLIGAPHYYEYQTRGQVSVCP 496	
QY	481 PRGORARWOCDAVLYGEOQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540	
DB	497 PRGORARWOCDAVLYGEOQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556	
QY	541 HGTSGSGISPSHSORISAGSKSLPRLOYFGOSLGGQDLTMDGLVDLTGAGQHVILLRSQ 600	
DB	557 HGTSGSGISPSHSORISAGSKSLPRLOYFGOSLGGQDLTMDGLVDLTGAGQHVILLRSQ 616	
QY	601 PVLVRKAIMFNPREVARNVFECDQVWKGAGEVRVCLHVQKSTRDLREGQIQSVVT 660	
DB	617 PVLVRKAIMFNPREVARNVFECDQVWKGAGEVRVCLHVQKSTRDLREGQIQSVVT 676	
QY	661 YDLALDSGRPHSRVFNENKSTRTOTVGLTQTCETLKLQIPNCIEDPVSPIVLRINF 720	
DB	677 YDLALDSGRPHSRVFNENKSTRTOTVGLTQTCETLKLQIPNCIEDPVSPIVLRINF 736	
QY	721 SLVGTPLSAFNLPRVLAEDAQLFTALPFFKFNCCNDNICDDLSITFSFMSLDCLVVG 780	
DB	737 SLVGTPLSAFNLPRVLAEDAQLFTALPFFKFNCCNDNICDDLSITFSFMSLDCLVVG 796	
QY	781 GPREFNVTVTRNDGDSYRTQVTFPPDLDSVRKYSTIQNRSQSRWSRLACESASSTEV 840	
DB	797 GPREFNVTVTRNDGDSYRTQVTFPPDLDSVRKYSTIQNRSQSRWSRLACESASSTEV 856	
QY	841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANTYSENMPRTNKTEF 900	
DB	857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANTYSENMPRTNKTEF 916	
QY	901 QLELPVKYAVYVMTVSHGYSVKYLNFTASENTSRVMQHOVQVSNLQORSPLTSLVLPV 960	
DB	917 QLELPVKYAVYVMTVSHGYSVKYLNFTASENTSRVMQHOVQVSNLQORSPLTSLVLPV 976	
QY	961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020	
DB	977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1036	
QY	1021 FFGIOBEFNATLKGNLSPFDWYIKTSHNHLIIYSTABILENDSVFTLLPGOGAFVRSQTET 1080	
DB	1037 FFGIOBEFNATLKGNLSPFDWYIKTSHNHLIIYSTABILENDSVFTLLPGOGAFVRSQTET 1096	
QY	1081 KVEPEFVNPPLPIVGVSSVGGILLALALITAAALYKLGFFRQYKDMMSSEGGPPGAEPO 1137	
DB	1097 KVEPEFVNPPLPIVGVSSVGGILLALALITAAALYKLGFFRQYKDMMSSEGGPPGAEPO 1153	
RESULT 5		
AA014428		
ID	AA014428 standard; protein; 1153 AA.	
XX	AA014428;	
AC	AA014428;	
XX	03-MAY-2002 (first entry)	
DT	Integrin Mac-1 alpha subunit.	
XX		
DE	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
XX	open conformation; integrin related inflammatory disorder;	
KW	integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW	reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW	viral infection; cancer; gene therapy; vaccine;	
KW	bioactive agent screening.	
XX	Unidentified.	
OS		
XX	WO200204521-A2.	
PN	17-JAN-2002.	
XX		
PD	09-JUL-2001; 2001WO-US021805.	
XX	07-JUL-2000; 2000US-0216600P.	
XX	(CALY) CALIFORNIA INST OF TECHNOLOGY.	
PA	(BLOO-) CENT BLOOD RES.	
XX	Springer T;	
PI	WPI; 2002-148167/19.	
XX	New integrin I domain protein having alteration in at least 2	
DR	noncontiguous regions and exits in an open conformation, useful for	
XX	treating, preventing or suppressing inflammatory or immunological	
PT	disorders.	
PT	Example 1; Fig 1F; 90pp; English.	
XX	The invention comprises structurally biased variant integrin inserted (1)	
CC	domain proteins, wherein the alterations to the protein occur in at least	
CC	two noncontiguous regions. Specifically the variant integrin I domain	
CC	proteins are structurally biased to exist in the open conformation,	
CC	thereby altering the binding ability of the protein. The invention also	
CC	comprises nucleic acids encoding the variant integrin I domain proteins.	
CC	The integrin I domain proteins and nucleic acids are useful for treating,	
CC	preventing or suppressing integrin related inflammatory and immunological	
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain	
CC	proteins and nucleic acids can also be used for treating: ischaemia/	
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral	
CC	infection; and cancer. The variant integrin I domain nucleic acids and	
CC	proteins may be used in gene therapy, as vaccines and to screen for	
CC	bioactive agents. The present amino acid sequence represents the Mac-1	
CC	alpha subunit of integrin	
XX	Sequence 1153 AA;	
SQ		
Query Match	99.9%; Score 5868; DB 5; Length 1153;	
Best Local Similarity	99.6%; Pred. No. 0;	
Matches 1133; Conservative	3; Mismatches 1; Indels 0; Gaps 0;	

pharmaceutical composition comprising the binding domain-immunoglobulin fusion protein or polynucleotide and a carrier, and treating a subject having or suspected of having a malignant condition or a B-cell disorder. The binding domain-immunoglobulin fusion protein is useful for treating a subject having or suspected of having a malignant condition or a B-cell disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis, myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple sclerosis or autoimmune disease. The present sequence is a binding domain -immunoglobulin fusion protein-associated protein sequence. Note: The sequence data for this patent formed part of the printed specification and is also available in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not identified the sequences in the printed specification by their SEQ ID number therefore none of the sequences can be explicitly identified.

XX Sequence 1153 AA;

Query Match 99.9%; Score 5968; DB 7; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQK 136
QY 121 FPALRGCPQEDSDIAPLIDGSGSIIPHPDFRMKEFVSTVMEQKKSKTLFSLMQYSEEF 180
DB 137 FPALRGCPQEDSDIAPLIDGSGSIIPHPDFRMKEFVSTVMEQKKSKTLFSLMQYSEEF 196
QY 181 RIHFTKEFONNPNRSLKPIQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
DB 197 RIHFTKEFONNPNRSLKPIQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 256
QY 241 TDEKEFGDPLGYEDVPEADREGVIRVIGVGAFAFRSEKSRQELNTIASKPPRDHVFQVN 300
DB 257 TDEKEFGDPLGYEDVPEADREGVIRVIGVGAFAFRSEKSRQELNTIASKPPRDHVFQVN 316
QY 301 NFPAKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAAITSNGLLSTVGSYDWAG 360
DB 317 NFPAKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAAITSNGLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGGQVSCPL 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGGQVSCPL 496
QY 481 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRQQRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLRLRSQ 600
DB 557 HGTSGSGISPHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLRLRSQ 616
QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
DB 617 PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSAFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSIVILRLNF 720
DB 677 YDLALDSGRPHSAFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSIVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPEKNCNDNICQDDLSITFSFMSLDCILVWG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPEKNCNDNICQDDLSITFSFMSLDCILVWG 796

QY 781 GPREENVTVRNDGEDSYRTQVTFPPDLDSYRKUSTLQNSORSQSWRLACESASSTEV 840
DB 797 GPREENVTVRNDGEDSYRTQVTFPPDLDSYRKUSTLQNSORSQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVVTSHGVSTKYLVNTASNTSRVMOHQYQVSNLQORSPLISLVFLVPV 960
DB 917 QLELPVKYAVYVVTSHGVSTKYLVNTASNTSRVMOHQYQVSNLQORSPLISLVFLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDVSFTLLPGQGAFFVRSOTET 1080
DB 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDVSFTLLPGQGAFFVRSOTET 1096
QY 1081 KVTPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1137
DB 1097 KVTPFEVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSSEGGPPGAEPQ 1153
RESULT 7
ID AAR04136 standard; protein; 1153 AA.
XX AAR04136;
XX AC AC
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 07-SEP-1990 (first entry)
XX
DE Alpha subunit of Mac-1 leukocyte adhesion receptor.
XX
KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;
KW non-specific defence system; integrin gene superfamily.
XX
OS Synthetic.
FH Key
FH Region 1..116
FT /label= signal_peptide
FT Modified-site 86..88
FT /label= putative N-glycosylation site
FT Modified-site 240..242
FT /label= putative N-glycosylation site
FT Modified-site 391..393
FT /label= putative N-glycosylation site
FT Modified-site 469..471
FT /label= putative N-glycosylation site
FT Modified-site 693..695
FT /label= putative N-glycosylation site
FT Modified-site 697..699
FT /label= putative N-glycosylation site
FT Modified-site 735..737
FT /label= putative N-glycosylation site
FT Modified-site 802..804
FT /label= putative N-glycosylation site
FT Modified-site 881..883
FT /label= putative N-glycosylation site
FT Modified-site 901..903
FT /label= putative N-glycosylation site
FT Modified-site 912..914
FT /label= putative N-glycosylation site
FT Modified-site 941..943
FT /label= putative N-glycosylation site
FT Modified-site 947..949
FT /label= putative N-glycosylation site
FT Modified-site 979..981
FT /label= putative N-glycosylation site

FT Modified-site 994..996
 FT /label= putative N-glycosylation site
 FT 1022..1024
 FT /label= putative N-glycosylation site
 FT 1045..1047
 FT /label= putative N-glycosylation site
 FT 1051..1053
 FT /label= putative N-glycosylation site
 FT 1076..1078
 FT /label= putative N-glycosylation site
 FT 1106..1134
 FT /label= putative_transmembrane_region
 XX
 PN EP364690-A.
 XX
 XX 25-APR-1990.
 XX
 XX 17-AUG-1989; 89EP-00115159.
 XX
 XX 23-AUG-1988; 88US-00235353.
 PR 09-MAR-1989; 89US-00321239.
 PR
 XX (DAND) DANA FARBER CANCER INST INC.
 FA
 XX Springer TA, Corbi A;
 PI
 XX WPI; 1990-125938/17.
 DR N-PDSB; AAQ04043.
 DR
 XX
 FT New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating
 FT inflammation and viral infections, and in diagnosis.
 FT
 XX Disclosure; Page ?; 3pp; English.
 XX
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.
 CC recognition of and migration to sites of inflammation. It also attaches
 CC to cellular substrates as part of this function making it useful in
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25
 CC -MAR-2003 to correct PA field.)
 CC
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key
 CC and pages
 CC
 XX Sequence 1153 AA;
 SQ
 Query Match 99.7%; Score 5858; DB 2; Length 1153;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 1132; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ENLDTENAMTQENARGFGQSVQVQGGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 60
 DB 17 ENLDTENAMTQENARGFGQSVQVQGGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 76
 QY 61 RLQVPVAVNMVSLGLSLAATSPQQLACGPTVHQTCSNTYVKGCLFLGSLNRQQPQK 120
 DB 77 RLQVPVAVNMVSLGLSLAATSPQQLACGPTVHQTCSNTYVKGCLFLGSLNRQQPQK 136
 QY 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 180
 DB 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196
 QY 181 RIHFTFEFQNNPRSLVKPITQLGTHTHATGVRKVIRELNTNGARKNAKILIVI 240
 DB 197 RIHFTFEFQNNPRSLVKPITQLGTHTHATGVRKVIRELNTNGARKNAKILIVI 256
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRHHVQVN 300
 DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRHHVQVN 316
 QY 301 NFEALKTTQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATNSGPIITVGSYDWAG 360
 DB 317 NFEALKTTQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATNSGPIITVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 420
 DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGVAAAIIILNRVQSLVIGAPRYQHIGLVAMFR 436
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEYQTRGGQVSCPL 480
 DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVIGAPHYEYQTRGGQVSCPL 496
 QY 481 PRGQARWQCDADVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGABEDNRGAYLIF 540
 DB 497 PRGQARWQCDADVLYGEGQGPWGRFGAALTIVLGVNVDGKLTVDVAIGABEDNRGAYLIF 556
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTWAGHVVLLRSQ 600
 DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTWAGHVVLLRSQ 616
 QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVQKSTDRDLREGQIQSVVT 660
 DB 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVQKSTDRDLREGQIQSVVT 676
 QY 661 YDLALDSGRPHSRAVFNSTKSTRTQVGLTCTCETLKLQLPNCIEDPSPVILRLNF 720
 DB 677 YDLALDSGRPHSRAVFNSTKSTRTQVGLTCTCETLKLQLPNCIEDPSPVILRLNF 736
 QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
 DB 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
 QY 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
 DB 857 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVPV 960
 DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVPV 976
 QY 961 RLNTQVTDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
 DB 977 RLNTQVTDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
 QY 1021 PFGIQEENATLKGNSLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFAVRSQET 1080
 DB 1037 PFGIQEENATLKGNSLSPDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGAFAVRSQET 1096
 QY 1081 KVEPFEPVNPPLTVGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137
 DB 1097 KVEPFEPVNPPLTVGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153
 RESULT 8
 ADM9589
 ID ADM9589 standard; protein; 1152 AA.
 XX
 AC ADM9589;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Human integrin alphaM subunit precursor protein.
 XX
 KW integrin alpha subunit; beta; antiposrotic; thrombolytic; anticoagulant;
 KW osteopathic; cytosolic; immunosuppressive; anti-inflammatory;
 KW neuroprotective; antisticking; immunotherapy; inflammatory;
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;
 KW alpham.
 XX
 OS Homo sapiens.
 XX

Key	Location/Qualifiers	
FT	Misc-difference 965	
FT	/note= "Encoded by CCC"	
XX		
PN	WO2004007530-A2.	
PD	22-JAN-2004.	
XX		
XX	17-JUL-2003; 2003WO-US022301.	
XX	17-JUL-2002; 2002US-0396783P.	
PR	17-JUL-2002; 2002US-0396790P.	
PR	11-SEP-2002; 2002US-0410135P.	
XX		
PA	(BLOO-) CENT BLOOD RES INC.	
XX		
PI	Springer TA, Takagi J;	
XX	WPI; 2004-122877/12.	
DR	N-PSDB; ADM9598.	
XX		
PT	Novel modified integrin protein having extracellular domains of integrin alpha and beta subunits or integrin alpha1 and beta3 subunit, useful for treating integrin mediated disorders.	
XX		
PS	Disclosure; SEQ ID NO 4; 232pp; English.	
XX		
CC	The invention relates to a novel isolated or recombinant modified integrin protein having extracellular domains of integrin alpha and beta subunits where one of the subunits has one or more mutations, an altered surface feature or an amino acid substitution or internal deletion, extracellular domains of the integrin beta subunit that comprise a mutation that alters a non-cysteine residue to cysteine or extracellular domains of integrin alpha and beta subunits. The polypeptide of the invention demonstrates antipsoriatic, thrombolytic, anticoagulant, osteoprotective, cytostatic, immunosuppressive, antiinflammatory, neuroprotective and anticikling activities and may be useful for immunotherapy in order to prevent or treat an integrin-mediated disorder such as an inflammatory disorder, an autoimmune disorder, thrombosis, cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple sclerosis. The current sequence is that of the human integrin alpha subunit precursor protein of the invention.	
XX		
SQ	Sequence 1152 AA;	
Query Match		39.6%; Score 5852.5; DB 8; Length 1152;
Best Local Similarity		99.6%; Pred. No. 0;
Matches 1132; Conservative		3; Mismatches 1; Indels 1; Gaps 1;
QY	1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI	60
DB	17 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAANORGSLYQCDYSTGSCPEI	76
QY	61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGNSLNRQOPQK	120
DB	77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGNSLNRQOPQK	136
QY	121 FFEALRGCPQESDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTFLSLMQYSEEF	180
DB	137 FFEALRGCPQESDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTFLSLMQYSEEF	196
QY	181 RHFTKEFQNNPNRSLVKPIITQLLGRTHATGVRKPIRELLNITNGARKNAFKILVI	240
DB	197 RHFTKEFQNNPNRSLVKPIITQLLGRTHATGVRKPIRELLNITNGARKNAFKILVI	256
QY	241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPRPDHVFQYN	300
DB	257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPRPDHVFQYN	316
QY	301 NFALKTIONQLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG	360
DB	317 NFALKTIONQLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG	376
QY	361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRNVOSLVLGAPRYOHTGLVAMER	420
DB	377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRNVOSLVLGAPRYOHTGLVAMER	436
QY	421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSCPL	480
DB	437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSCPL	496
QY	481 PRGQARWOCDAVLYGEOGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAYLFL	540
DB	497 PRG-RARWOCDAVLYGEOGQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAYLFL	555
QY	541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ	600
DB	556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDTVGAQGHVLLRSQ	615
QY	601 PVLVRKAIIMEFNPREVARNVFECDQVKGKAGEVRVCLHVQKSTRDRREGQIQSVVT	660
DB	616 PVLVRKAIIMEFNPREVARNVFECDQVKGKAGEVRVCLHVQKSTRDRREGQIQSVVT	675
QY	661 YDLALDSGRPHSRAPNETKNSTRQTVGLTQTCETILKQLPNCIEDPVPSPVILRLNF	720
DB	676 YDLALDSGRPHSRAPNETKNSTRQTVGLTQTCETILKQLPNCIEDPVPSPVILRLNF	735
QY	721 SLVGTPLSAFGNLRPVLAEDAORLFTALPFEKNCGNDNICQDDLSTITFSMELDCLVVG	780
DB	736 SLVGTPLSAFGNLRPVLAEDAORLFTALPFEKNCGNDNICQDDLSTITFSMELDCLVVG	795
QY	781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSLACESASSTEV	840
DB	796 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSLACESASSTEV	855
QY	841 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGKLLKLLKANTSENMPRTNKTEF	900
DB	856 SGALKSTCSINHPIPPENSEVTNITFDVDSKASLGKLLKLLKANTSENMPRTNKTEF	915
QY	901 QLELPVKYAVYVMVTSHGVSSTYKLNFTASNTSRVMQHOYQVSNLQORSIPISLVLVPV	960
DB	916 QLELPVKYAVYVMVTSHGVSSTYKLNFTASNTSRVMQHOYQVSNLQORSIPISLVLVPV	975
QY	961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP	1020
DB	976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORICDIP	1035
QY	1021 FFGIOEEFNATLKGNSLSPDWYIKTSHNLLIIVSTABILPNDVSTLLPGQGFVRSQTET	1080
DB	1036 FFGIOEEFNATLKGNSLSPDWYIKTSHNLLIIVSTABILPNDVSTLLPGQGFVRSQTET	1095
QY	1081 KVEPPEVPNPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEQ	1137
DB	1096 KVEPPEVPNPLPIVGVSSVGGLLLLALITAALYKLGFFKQYKDMMSGGPPGAEQ	1152
RESULT 9		
ADP12435		
ID	ADP12435 standard; protein; 1152 AA.	
XX		
AC	ADP12435;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Protein encoded by mRNA of the invention #45.	
XX		
KW	transplant rejection; immune system; rheumatoid arthritis; lupus;	
XX	inflammatory bowel disease; multiple sclerosis; HIV; AIDS.	
OS	Homo sapiens.	
XX		
PN	WO2004042346-A2.	
XX		
PD	21-MAY-2004.	
XX		
PF	24-APR-2003; 2003WO-US012946.	

XX 24-APR-2002; 2002US-00131831.
PR 20-DEC-2002; 2002US-00325899.
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
PA Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX WPI; 2004-400724/37.
DR
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
PS Claim 65; SEQ ID NO 2444; 1762pp; English.
XX
CC The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC protein that is encoded by the mRNA of the invention.
XX
SQ Sequence 1152 AA;

Query Match 99.5%; Score 5845.5; DB 8; Length 1152;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQIIVANORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQIIVANORGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLACGTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPQLLACGTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQSDIAFLIDGSGIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 180
DB 137 FPEALRGCPQSDIAFLIDGSGIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196

QY 181 RIHETFEFQNNPNSRIVKEITOLLGRTHATGVRKVIKELNITNGARKNAFKILIVI 240
DB 197 RIHETFEFQNNPNSRIVKEITOLLGRTHATGVRKVIKELNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAPRSEKSRQELNTIASKPPRPHVQVN 300
DB 257 TDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAPRSEKSRQELNTIASKPPRPHVQVN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPFHEMSQGFSAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSPFHEMSQGFSAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFNIMTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMER 420
DB 377 GVFLYTSKEKSTFNIMTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMER 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYIEYQTRGGQSVYCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYIEYQTRGGQSVYCP 496

QY 481 PRGORARQCDVAVLYGEGQGWFRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 540
DB 497 PRG-RARWQCDVAVLYGEGQGWFRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 555

QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQHVLLLRQ 600
DB 556 HGISGSGISPSHSORITAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQHVLLLRQ 615

QY 601 PVLIRVKAIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVKQKSTRDLREGQIQSVVT 660
DB 616 PVLIRVKAIMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVKQKSTRDLREGQIQSVVT 675

QY 661 YDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVSPIVLINP 720
DB 676 YDLALDSGRPHSRVAFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVSPIVLINP 735

QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDDLSITTFGMSLDCIVVG 780
DB 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKNCGNDNICQDDLSITTFGMSLDCIVVG 795

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 840
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 855

QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 915

QY 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 960
DB 916 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 975

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035

QY 1021 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVETILLPGQAFVRSQDET 1080
DB 1036 FFGIOEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVETILLPGQAFVRSQDET 1095

QY 1081 KVPFPFVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137
DB 1096 KVPFPFVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1152

RESULT 10
ADP44061
ID ADP44061 standard; protein; 1163 AA.
XX
AC ADP44061;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human CD11C protein SEQ ID NO:14.
XX
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.
XX
OS Homo sapiens.
XX
PN W02004053094-A2.
XX
PD 24-JUN-2004.
XX
PF 08-DEC-2003; 2003WO-US039208.
XX
PR 06-DEC-2002; 2002US-0431522P.
XX
PA (PPDP-) PPD DEV LP.
XX
PI Dunn SJ, Holzmayer TA;
XX
PI WPI; 2004-480928/45.
DR
DR N-FSDB; ADP44060.
XX
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

PT	agent for IBD.	
XX		
PS	Example; SEQ ID NO 27; 151pp; Japanese.	
XX		
CC	The invention comprises marker genes for inflammatory bowel disease (IBD)	
CC	- FcgammaR IIa, FcgammaR IIb, Mig, NKG-2, hexokinase, HM74, REG III,	
CC	LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.	
CC	The DNA and protein sequences of the invention are useful in the	
CC	detection and treatment of IBD (e.g. ulcerative colitis and Crohn's	
CC	disease). The present amino acid sequence represents an IBD marker	
CC	protein of the invention.	
XX		
XX	Sequence 1163 AA;	
XX		
XX	Query Match	
XX	Best Local Similarity 59.1%; Score 3473; DB 8; Length 1163;	
XX	Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;	
QY	1 FNLDTENAMTFOENARGFGQSVWOLQGSRRVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 60	
DB	20 FNLDETELTPRVDAGFGDSVQVANSVVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79	
QY	61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGLCLFLFSNLRQOPQK 120	
DB	80 GLQVPPRAVNMVSLGLSLASTTSPSOLLACGPTVHHECGRNMYLTGLCPFLGPT--QLTQR 137	
QY	121 PPEALRGCPQSDSDIAFLIDGSGSII PHDFRMKEFVSTVMEOLKSKTLESLMYSSEF 180	
DB	138 LPVSRQECPEQDIFVLIDGSGSISRRNFATMNFVRAVTSQFQRPSTQSLMQFSNKF 197	
QY	181 RIHTEFEFONNPRSLVKPTLTOLLGRTHTATGVKRVIRELLNITNGARKNAKILIVI 240	
DB	198 QTHFTFEFRITSNPLSLASVHQLQGFYITATQNVVHRLPHASYGARRDAIKILIVI 257	
QY	241 TDGEKFGDGLYEDVIPADREGVIRYVIGVDAPRSEKSRQELMTIASKPPRDRHVQVN 300	
DB	258 TDGKKEGSLDYKDVIPWADAAGIIRYAIQVGLAFQNRNSWKNELNDIAKPSQEHIFKE 317	
QY	301 NFEALKITQNLREKIFAIEGTQTSSSSFEHENGQEGSAITNSGPLLSTVGSYDWAG 360	
DB	318 DFDALKIQNLQKIFAIEGTETSSSFELEMAQEGFSAVTPDGVGLVAVGVSFTWSG 377	
QY	361 GVFLTSKEKSTFINMTVDSDMDNDAYLGYAAAIIILNRVQSLVLCAPRYCHIGLVAMFR 420	
DB	378 GAFLYPPNMSPTFINNSGVNDVMSYLGYSYELALWKGVQSLVLCAPRYQHTGKAVFT 437	
QY	421 QNTGMWESNANVKGFTQICAYFGASLCSDVDVSDNGSTDLVLCAPRYHQTGQGVSVCP 480	
DB	438 QVSRQWRMKAETVGTQIGSYFGASLCSDVDVSDNGSTDLVLCAPRYHQTGQGVSVCP 497	
QY	481 PRQPARQCDVLYGEGQPMGRFGAALTVLGVNGDKLTVDVAIGAPCEEDNRCVAVLF 540	
DB	498 PRGWR-RWMCDAVLYGEGQPMGRFGAALTVLGVNGDKLTVDVIGAPCEENRCVAVLF 556	
QY	541 HGTSGSISPHSRIAGSKLSPRIQYFGQSLSGGQDLTMDCGLVDLTGAQGHVLLRSQ 600	
DB	557 HCVLGPSISPHSRIAGSKLSRLQYFGQALSGGQDLTQDGLVDLVAGARQVLLRTR 616	
QY	601 PVLRYKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRLRGQTSVVT 660	
DB	617 PVLWVGVSQFTPAEIPRSAPFCEQVWSEQLTVQSNICLYIDKRSKLLGSRDLQSSVT 676	
QY	661 YDLALDSGPHSRVAFNETKSTRQTVLGTQTCETLKLQCLPNCIEPVPVILRLNF 720	
DB	677 LDALDPGLSPRATFOETKNSLSRVRVILGKAHCENFLLPSCVDSVPTITRLNF 736	
QY	721 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFPFKNCGNDNICQDLSITFSPMSLCLVVG 780	
DB	737 TLVGKELLAFNLRLPMAALARYFTASLPFEKNCQADHICQNDLIGISFSPGLKSLVG 796	
QY	781 GRPEFNVTVRMDGDSSTQVTFPPFLDLSYKRVSTLQNSQBSWLACASSTEV 840	
DB	797 SNLELNAEVWMDGSDSYGTTITTFSHPAGLSYRYVAEGQKQGLRSLHLCDSAPVG-- 854	
QY	841 SGALKSTSCSINHPIPIPPENSEVITFNITFDVDSKASLGNKLLKXANVTSENNMPTNKTEF 900	
DB	855 SQGTWSTSCRNHLIFRGGAQITFLATFDVSPKRAVLGDRLLLTANVSSENTPRTSKTTF 914	
QY	901 QLELPVKYAVYVMVTVSHGVSTKYLNFAS-ENTSRVMQHOYQVSNLQORSIPISLVLVP 959	
DB	915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974	
QY	960 VRLNQTIVMDRPOVTFSENLSSTCHTERLPSSDFLAELRKAPVNVNCSIAVCORICODI 1019	
DB	975 VELNQEAVMWDMVEVSHPPQNSLRCSSEKIAPPASDFLAHQKPNVLDSCSIAGCLFRCDV 1034	
QY	1020 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTABILFNDVSFTLLPGGAFVRQOTE 1079	
DB	1035 PSFSVQSELDFTLKGNSLFGWVRQILQKVSWSVABITFTSVYSQLPQGEAFMRAQT 1094	
QY	1080 TKVEPFVFNPLPLIVGSSVGGLLLLALITLAALYKLGFFKXQYKDMSE 1128	
DB	1095 TVLEKYVHNPTPLIVGSSIGGLLLALITAVLYKVGFFKXQYKEMMBE 1143	
RESULT 12		
ADQ17510		
ID	ADQ17510 standard; protein; 1163 AA.	
XX		
AC	ADQ17510;	
XX		
DI	26-AUG-2004 (first entry)	
XX		
DE	Human soft tissue sarcoma-upregulated protein - SEQ ID 327.	
XX		
KW	soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.	
XX		
OS	Homo sapiens.	
XX		
FN	WO2004048938-A2.	
XX		
PD	10-JUN-2004.	
XX		
PF	26-NOV-2003; 2003WO-US038193.	
XX		
PR	26-NOV-2002; 2002US-0429739P.	
PA	(PROT-) PROTEIN DESIGN LABS INC.	
XX		
PI	Aziz N, Ginsburg WM, Zlotnik A;	
XX		
DR	WPI; 2004-441208/41.	
XX		
PT	Early detection of soft tissue sarcoma comprises determining expression	
PT	of a gene in a first soft tissue sample and a normal soft tissue sample	
PT	and comparing the gene expression, also useful in treating soft tissue	
PT	sarcoma.	
XX		
PS	Example 2; SEQ ID NO 327; 210pp; English.	
XX		
CC	The invention relates to a novel method for detecting soft tissue sarcoma	
CC	which comprises obtaining a first soft tissue sample from an individual	
CC	and a normal soft tissue sample from the same or different individual,	
CC	determining the expression of a gene in both samples and comparing the	
CC	expression of the gene in both soft tissue samples, where a higher level	
CC	of protein expression in the first soft tissue sample indicates the	
CC	presence of soft tissue sarcoma. The method of the invention has	
CC	cytostatic applications and may be useful for detecting soft tissue	
CC	sarcoma, possibly via gene therapy or vaccine production. The nucleic	
CC	acid sequences may be useful in diagnostic and screening applications.	
CC	The current sequence is that of a human soft tissue sarcoma-upregulated	
CC	protein of the invention. The current sequence is not shown within the	
CC	specification per se but was submitted in CD format by the inventor.	
XX		
SQ	Sequence 1163 AA;	

Query Match		59.1%; Score 3473; DB 8; Length 1163;
Best Local Similarity		61.3%; Pred. No. 6.9e-280;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;		
Qy	1	FNLDTENAMTFQENARGFQGVVQLQGSFVVGAPQEIIVAAQNRSLVQCDYSTGSCBPI 60
Db	20	FNLDTEELTAFRVDSAGFDGSVVQVANSVVVWVWVAPQKTAANQGTGLYQCGYSTGACEPI 79
Qy	61	RLOVPVEAVNMISGLSLAATTSFPLLACGPTVHQTCSNTYVVKGLCFGLFSGNLQOQPK 120
Db	80	GLQVPPEAVNMISGLSLATTSFPLLACGPTVHHECGENMYLTGLCFLLGPT--QLTQR 137
Qy	121	FPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKXSKTLFSLMOYSEEF 180
Db	138	LPVSRQECPRQEQDIFVLIDGSGSISSRNFAFMNFRAVISQFORPSTQFSLMQFSNKF 197
Qy	181	RIHFTFKFONNPRSLVKPIITOLGLRTHATGVKRVIRELLNITNGARKNAFKILLIVI 240
Db	198	QTHFTFEFRITSNPLSLASVHQLQGYTYTATAIQNVVHRLFHASYGARRDATKILIVI 257
Qy	241	TGGEKFGDPLGYEDVIEPADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
Db	258	TGKKEGDSLKYDQVIMADAAAGIIRYVIGLAFQNRNSWKELNDIASQSHIFKVE 317
Qy	301	NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db	318	DFDALKIQNLQKKEIFAIEGTETTSSTSSFELEMAQEGFSAVFTPDGVPVLAGVGSFTWSG 377
Qy	361	GVFLYTSKESKSTFLNMTRVSDMDNDAVLGYAAAIILNRVOSLVIGAPRYOHIGLVAMFR 420
Db	378	GAFLYPNMSPFTFNMQENVDMDRSLVGLSTELALMKGVQSLVIGAPRYOHTKAVIFT 437
Qy	421	QNTGMWESNANVKGTQIGFYAGSLCSVDVDSNGSTDLVLIGAPHYVEOTRGGQVSCPL 480
Db	438	QVSRQWRMKAETVGTQIGFYAGSLCSVDVDTGSDTLVLIGAPHYVEOTRGGQVSCPL 497
Qy	481	PRGARQWQDAVLYGQGPQWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAVYLF 540
Db	498	PRGWR-RWDCDAVLYGQGPQWGRFGAALTVLGVDNGDKLTDVVIGAPGEENRGAVYLF 556
Qy	541	HGTSGSGISPSHSORLAGSKLSPLOYFGSLSGQDLTWGDLVLTVGAGHVLRLRSQ 600
Db	557	HGVLGFSISPSHSORLAGSKLSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLIRTR 616
Qy	601	PVLVRKAIMEFNEPREVARNVFEQNDQVWKGKEAGEVRVCLHVOKSTEDRLREGIOSVVT 660
Db	617	PVLWVGVMQFIPAEIPRSFAFECREQVSEQTIVQSNICLIYIDKRSKNLLGSDLOSST 676
Qy	661	YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDVPSPVILRLNF 720
Db	677	LDLALDPGLRSPRATFQETKNRSLRVVLGKAHCENFNLLPSCVEDSVTPITLRLNF 736
Qy	721	SLVGTPLSAFGLNRPVLAEDAQRLETFALPFEKNCNDNICODDLITFSFMSLDCLVVG 780
Db	737	TLVGKPLLAFLNRPMLAALQRYFTASLPFEKNCADHICQDNILGIFSFGPKSLLVG 796
Qy	781	GPREFNVTVVRNMGDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db	797	SNLELNAEVMWMDGSDSYGTITTFHSHPAGLSRYRYAEGQKQQLRSLHLTCDSPVGV-- 854
Qy	841	SGALKSTSCSINHPFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db	855	SOGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVLGDRLTLTANVSSENTPRTSKTTF 914
Qy	901	QLELPVKYAVVMVTSYSHGVSTKYLNFTAS-ENTSRVMQHOYQVNSLQGRSLPISLVFLVP 959
Db	915	QLELPVKYAVITVVSHEQFTKLYNFSESEESHVAMERYQVNNLQGRDLPLVSNFWVP 974
Qy	960	VRNLQTVIWRDPVTFSESNLSSTCHTKERLPSHSDFLAELRKAQVNVCSIAVCORQCDI 1019
Db	975	VELNQAQVMDVEVSHVPQNPRLCSSEKAPASDFLAHIQRPVNDCCSIAGCLFRCDV 1034
Qy	1020	PFQIQEENATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db	1035	PSFSVQEBLDFTLKGNLSFGWVRQLQKKVSVVAEITFTDTSVYSLQFGQAFMRAQTT 1094
Qy	1080	TKVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKDMWSE 1128
Db	1095	TVLEKYKVHNTPLIVGSSIGGLLLALITAVLYKVGFFKROYKEMMEE 1143
RESULT 13		
ID	AA07120	
AC	AA07120 standard; protein; 1163 AA.	
XX	AA07120;	
XX	25-MAR-2003 (revised)	
DT	05-FEB-1991 (first entry)	
DE	p150.95 alpha subunit encoded by clone lambda47.	
XX	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;	
KW	rhinovirus.	
XX	Synthetic.	
XX	Key	Location/Qualifiers
XX	Region	1..19
FT	Region	/label= signal peptide
FT	Region	20..44
FT	Modified-site	/label= N-terminus
FT	Modified-site	61..63
FT	Modified-site	/label= glycosylation site
FT	Modified-site	89..91
FT	Modified-site	/label= glycosylation site
FT	Modified-site	385..387
FT	Modified-site	/label= glycosylation site
FT	Modified-site	392..394
FT	Modified-site	/label= glycosylation site
FT	Modified-site	697..699
FT	Modified-site	/label= glycosylation site
FT	Modified-site	735..737
FT	Modified-site	/label= glycosylation site
FT	Modified-site	899..901
FT	Modified-site	/label= glycosylation site
FT	Modified-site	904..906
FT	Modified-site	/label= glycosylation site
FT	Modified-site	939..941
FT	Modified-site	/label= glycosylation site
FT	Modified-site	1050..1052
FT	Modified-site	/label= glycosylation site
FT	Domain	1108..1133
FT		/label= transmembrane
XX	W09010646-A.	
XX	20-SEP-1990.	
XX	09-MAR-1990;	90WO-US001257.
XX	09-MAR-1990;	90WO-US001257.
XX	(DAND) DANA FARBER CANCER INST INC.	
XX	Corbi AA, Springer TA;	
XX	WPI; 1990-304985/40.	
XX	N-PSDB; AAQ06068.	
XX	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-	
XX	unit of p150.95 cell surface adhesion receptor, opt. together with a beta	
XX	chain of CD-18 family.	
XX	Disclosure; Fig 3; 59pp; English.	

CC Clone lambda X47 was isolated from a cDNA library constructed from total
CC RNA extracted from phorbol myristate acetate stimulated HL-60
CC myelomonocytic cells. The library was screened with oligonucleotide
CC probes based on tryptic peptide fragments of p150.95. The sequence can be
CC attached to appropriate control elements and expressed in prokaryotic and
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral
CC infection because it interacts with ICM-1 and inhibits cell-virus
CC attachment. It can also be used as an anti-inflammatory agent. See also
CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 1163 AA;

Query Match 59.0%; Score 3464; DB 2; Length 1163;
Best Local Similarity 61.1%; Pred. No. 3.9e-295;
Matches 690; Conservative 138; Mismatches 279; Indels 6; Gaps 4;

QY 1 ENLDTENAMTFQENARGFGQSVQVLOGSVVVGAPQEIIVAAANQSGLYQCDYSTGSCPEI 60
DB 20 FNLDETELTAFRVDSAGDGSVVQVANSVVVVGAPQKITAANQTGLYQCYSTGACEPI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLCFLFGSNLRQDPQK 120
DB 80 GLQVPPPEAVNMSLGLSLAATTSPQLLACGPTVHHECGRNMYLTGLCPLLGPT--QLTQR 137

QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKEFVSTVMEQLKSKTILSLMQYSEEF 180
DB 138 LPVSRQECPRQEDIVFLIDGSGISRRNFATMNFVRAVISQFQRPSTQFSLMOFSNKF 197

QY 181 RIHTEFEFQNNPRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAKILIVI 240
DB 198 QTHTEFEFRNTSPNLSLASVHQLOGTYTATQNVVHRLPHASYGARRDATKILIVI 257

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTIAKSPDRDHVQVN 300
DB 258 TDGKEGDSLDYKDVIPMDAAGIIRYALGVGLAFQNRNSWELNDIAKSPQEHIFKVE 317

QY 301 NFEALKTQNLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSVSGSYDNAG 360
DB 318 DFDALDKIQNLKKEIFAIEGTETTTSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHTGRAVFT 437

QY 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYYEOTRGQVSVCP 497

QY 481 PRGQARWOCDAVLGEGQCPWRFGAALTVLGVNGDKLTDVAIGARCEENRGAVILF 540
DB 498 PRGWR--RWCDAVLGEGQHPWRFGAALTVLGVNGDKLTDVIGAPGEENRGAVILF 556

QY 541 HGTSGSIGSPSHSQTAGSKLSPLOYFGQSLGGQDLMGDLVLTGVAQGHVLLRSQ 600
DB 557 HGVLPSPISPSHSQRTAGSRLQYFGQALSGGQDLTDGGLVDLAVGARQVILLRFR 616

QY 601 PVLRYKAIIMEFPREAVRVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLWGVSNQFIPAEIPRSFAFECREQVVEQTLVQSNICLYIDKRSKLLGSRDLQSSVT 676

QY 661 YDLALDSGPHRASVNETKSTRTQTVLGLTQTCETLKLQPNCEIDPVSPIVLRNPF 720
DB 677 LDIALDPGLSPRATFQETKNRSLSRVRLGKAHCENFNLLLPSCVEDSVTPITURLNF 736

QY 721 SLVGTPLSAFNLPRVLAEDAQRLETFALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 TLVKGKELLAFNRNPRMLAQAQYFTASIPPEKNCADHICQDNLGISFSFPGKSLVVG 796

QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACSSASTEV 840
DB 797 SNLELNAEVVMNDGEDSYGTITTFSPHAGLSYRYVAEGQKQGLRSLHLTCDSA--PVA 854

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLLKANTSENMMPTNKTEF 900
DB 855 SQGTWSTSCRINHLIFRGGAQITFLATFDVSPRAVLGDRILLTANVSSENNTPTSTKTF 914

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNTAS-ENTSRVMOHOYQVSNLQORSLSPLSLVLP 959
DB 915 QLELPVKYAVYVTVSSHEQTKYLNFSSEBESHVAMHRYQVNNJGQRLDPVSNINWVP 974

QY 960 VRLNQTVMIDRPOVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVCSIAVCQRIQCDI 1019
DB 975 VELNQEAVMWDMVEVSHQPNESLRCSSQKIAPPASDFLAHTQKNPVLDCSIAGCLRRPCDV 1034

QY 1020 PFGIQIEFNATLKNLSFDWYIKTSHNLLIYSTAEILNDSVFTLLPGQCAFVRQTE 1079
DB 1035 PFSFVQELDFTLLKGNLSPGWVRQILQKVSVSVSABITDTSVYSLPGQERFMRAQT 1094

QY 1080 TKVEPPEVNPPLPIVGVSSVGGILLALITAAALYKLGFEKRYKDMSE 1128
DB 1095 TVLEKYKVHNPPLPIVGVSSIGALLLALITAVLYKVGVFFKRYKEMME 1143

RESULT 14
AAW65091
ID AAW65091 standard; protein; 1163 AA.
XX
AC AAW65091;
XX
DT 28-SEP-1998 (first entry)
XX
DE Human Beta-integrin CD11c subunit protein.
XX
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;
KW rheumatoid arthritis.
XX
OS Homo sapiens.
XX
PN US5728533-A.
XX
PD 17-MAR-1998.
XX
PF 07-JUN-1995; 95US-00485618.
XX
PR 23-DEC-1993; 93US-00173497.
PR 05-AUG-1994; 94US-00286889.
PR 21-DEC-1994; 94US-00362652.
XX
PA (ICOS-) ICOS CORP.
XX
PI Van Der Vieren M, Gallatin WM;
XX
DR WPI; 1998-206565/18.
XX
PT Screening assay for modulators of integrin binding - using immobilised or
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.
XX
PS Example 5; Fig 1A-D; 106pp; English.
XX
CC This sequence represents a human beta-integrin CD11c subunit which is
CC used to describe a method for identifying compounds that modulate the
CC interaction of the beta-integrin alpha-d subunit with a binding partner
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha
CC -d binding partner, one of which is immobilised and the other of which is
CC labelled, in the presence of a test compound, and determining if the
CC compound affects binding between the alpha-d polypeptide and alpha-d
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment
CC comprising the cytoplasmic, transmembrane or extracellular domain of
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome
CC and rheumatoid arthritis
XX

SQ Sequence 1163 AA;

Query Match 58.7%; Score 3450; DB 2; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.7e-278;
Matches 689; Conservative 144; Mismatch 200;

[illegible]

Qy	1020	PFCIGI OEEFNATUKGNLSFDWYIKTSHNHLIIVSTAEILFNDSVFLLPGQAFVRSQTE	1079
Db	1035	PSRSGQEELFTLTKGNLSFGWRQIIKQKVSVVSAEIIIFDTSVGSQLPQGEAFMBAQTI	1094
Qy	1080	TKUPEPEVPNPPLIVGSSVGGLLLALITAALYKLGFPRKYOKNMSE	1128
Db	1095	TVLKRYKHNPPIIVGSSIGGLLLALITAVLYKVGFFRKYYKENMEE	1143

RESULT 15

RESULTS 13
AAB07361
ID AAB07361 standard; protein; 1163 AA.

AC AAB07361:

DT 17-JAN-2001 (first entry)

Human CD11c protein sequence.

Human; macrophage infiltration inhibition; alpha d integrin; leukocyte integrin; leu-CAM; leukointegrin; immune response; inflammation; leukocyte adhesion deficiency; LAD; Type 1 diabetes mellitus; atherosclerosis; multiple sclerosis; asthma; psoriasis; beta lung inflammation; acute respiratory distress syndrome; Crohn's disease; rheumatoid arthritis; central nervous system injury; CD1c.

OS Homo sapiens.

PN WO200029446-A1.

PD 25-MAY-2000.

PF 16-NOV-1999; 99WO-US027139

AA
PR 16-NOV-1998: 99ITS-00793043

PR 08-JUL-1999; 99US-00350259.
XX

PA
yy
(ICOS-) ICOS CORP.

PI Gallatin MW, Van Der Vieren M;

DR WPI; 2000-387751/33.

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.

Example 5; Fig 1; 270pp: English.

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, *Leu-CAMs*, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAAG60014 and BAB07359). The present sequence has approximately 56% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease

Sequence 1163 AA:

Query Match 58.7%; Score 3450; DB 3; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.7e-278;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds
(without alignments)
4119.157 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPGFAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5875	100.0	1137	10	US-09-902-481A-6
2	5868	99.9	1153	9	US-09-350-259-3
3	5868	99.9	1153	9	US-09-902-481A-1
4	5868	99.9	1153	14	US-09-891-943-3
5	5868	99.9	1153	14	US-10-144-259-30
6	5868	99.9	1153	14	US-10-207-655-176
7	5855	99.7	1137	10	US-09-902-481A-5
8	5852.5	99.6	1152	9	US-09-945-265-4
9	5845	99.5	1137	10	US-09-902-481A-4
10	5832	99.3	1137	10	US-09-902-481A-3
11	3473	59.1	1163	14	US-10-116-275-204
12	3450	58.7	1163	9	US-09-350-259-4
13	3450	58.7	1163	10	US-09-891-943-4

14	3417	58.2	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3417	58.2	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3401.5	57.9	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3401.5	57.9	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3239.5	55.1	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3239.5	55.1	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3228	54.9	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3228	54.9	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3218.5	54.8	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3218.5	54.8	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3207	54.6	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3207	54.6	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.5	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1552.5	26.4	1170	15	US-09-945-265-2	Sequence 2, Appli
28	1548.5	26.4	1223	16	US-10-408-765A-295	Sequence 295, App
29	1525	26.0	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
30	1350.5	23.0	494	9	US-09-350-259-103	Sequence 103, App
31	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
32	1229.5	20.9	413	9	US-09-350-259-101	Sequence 101, App
33	1157.5	19.7	413	10	US-09-891-943-101	Sequence 101, App
34	1157.5	19.7	413	9	US-10-177-550-2	Sequence 2, Appli
35	1149	19.6	1179	14	US-10-173-551-2	Sequence 103, App
36	1149	19.6	1179	14	US-09-984-130-103	Sequence 103, App
37	1102.5	18.8	1151	10	US-09-836-353A-103	Sequence 250, App
38	1102.5	18.8	1151	10	US-09-918-715-250	Sequence 810, App
39	1102.5	18.8	1179	17	US-10-474-794-250	Sequence 338, App
40	1102.5	18.8	1179	17	US-10-291-265-810	Sequence 35, Appl
41	1101	18.7	1188	14	US-10-291-265-810	Sequence 35, Appl
42	1097	18.7	1188	14	US-10-291-265-338	Sequence 4, Appli
43	1093.5	18.6	1189	10	US-09-984-130-35	
44	1093.5	18.6	1189	10	US-09-836-353A-35	
45	1093.5	18.6	1189	15	US-10-262-839-4	

ALIGNMENTS

RESULT 1

US-09-902-481A-6
; Sequence 6, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RTFRMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-6

Query Match	100.0%	Score	5875;	DB	10;	Length	1137;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches 1137;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
0;							
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLVCCDYSTGSCPEI	60				
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLVCCDYSTGSCPEI	60				
QY	61	RLOVPVEAVNMSLGLSLAATTSPPOLIACGPTVHQTCSNTYVKGKLCFLFGSNLRQQPQK	120				
Db	61	RLOVPVEAVNMSLGLSLAATTSPPOLIACGPTVHQTCSNTYVKGKLCFLFGSNLRQQPQK	120				

; Patent No. US20020062008A1		; GENERAL INFORMATION:	
; APPLICANT: Gallatin, Michael W.		; TITLE OF INVENTION: No. US20020062008A1el Human 2	
; APPLICANT: Van der Vieren, Monica		; FILE REFERENCE: 27866/35004	
; CURRENT APPLICATION NUMBER: 09/350,259		; CURRENT FILING DATE: 1999-07-08	
; EARLIER APPLICATION NUMBER: 09/193,043		; EARLIER FILING DATE: 1998-11-16	
; EARLIER APPLICATION NUMBER: 08/173,497		; EARLIER FILING DATE: 1993-12-23	
; EARLIER APPLICATION NUMBER: 08/286,889		; EARLIER FILING DATE: 1994-08-05	
; EARLIER APPLICATION NUMBER: 08/362,652		; EARLIER FILING DATE: 1994-12-21	
; EARLIER APPLICATION NUMBER: 08/943,363		; EARLIER FILING DATE: 1997-10-03	
; NUMBER OF SEQ ID NOS: 114		; SOFTWARE: PatentIn Ver. 2.0	
; SEQ ID NO 3		; LENGTH: 1153	
; TYPE: prf		; ORGANISM: Homo sapiens	
; US-09-350-259-3		; Query Match 99.9%; Score 5868; DB 9; Length 1153;	
		; Best Local Similarity 99.6%; Pred. No. 0;	
		; Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTTOENARGFGQSVVQOGSRVVVGAPQEIIVAAQNRGSLVQCDYSTGSCBPI	60
Db	17	FNLDTENAMTTOENARGFGQSVVQOGSRVVVGAPQEIIVAAQNRGSLVQCDYSTGSCBPI	76
QY	61	RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLCFGSLNLRQPOK	120
Db	77	RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLCFGSLNLRQPOK	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF	180
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF	196
QY	181	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIKRELIININGARKNAFKILIVI	240
Db	197	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIKRELIININGARKNAFKILIVI	256
QY	241	TGGEKFGDPLGYEDVITPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPDRHVFQVN	300
Db	257	TGGEKFGDPLGYEDVITPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPDRHVFQVN	316
QY	301	NFEALKTIONQLEKIFAIEGTGTGSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG	360
Db	317	NFEALKTIONQLEKIFAIEGTGTGSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL	480
Db	437	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL	496
QY	481	PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF	540
Db	497	PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVILF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ	600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLLRSQ	616
QY	601	PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	617	PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	676

QY 661 YDLALDSGRPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSQSWLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSQSWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVMVVTSYKYNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVMVVTSYKYNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDIP 1020
DB 977 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDIP 1036
QY 1021 FFGIQEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQET 1080
DB 1037 FFGIQEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQET 1096
QY 1081 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1153

RESULT 3

US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902, 481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match
Best Local Similarity 99.9%; Score 5868; DB 10; Length 1153;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVQVLOGSRVVVGAQOEIVAAANQORSLYQCDYSTGSCBPI 60
DB 17 FNLDTENAMTFOENARGFGQSVQVLOGSRVVVGAQOEIVAAANQORSLYQCDYSTGSCBPI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCLFLFSGNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCLFLFSGNLRQOPQK 136

RESULT 4

US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US2003007728A1

QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTWEOLEKSKTLLFSIMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTWEOLEKSKTLLFSIMQYSEEF 196
QY 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIRELNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIRELNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDARSEKSRQBELNTIASKPPDRHVFQV 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDARSEKSRQBELNTIASKPPDRHVFQV 316
QY 301 NFEALKTIONQUREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGPELLSTVGSVDNAG 360
DB 317 NFEALKTIONQUREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGPELLSTVGSVDNAG 376
QY 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLIGAPHYEQTREGGVSVCP 480
DB 437 QNTGWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLIGAPHYEQTREGGVSVCP 496
QY 481 PRGORARWQCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGORARWQCDAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAGHVLLRSQ 616
QY 601 PVLKVKALMEFNPREVARNVFECDQVVKGEAGEVRLVHVKQSTDRDLRREGQIQSVVT 660
DB 617 PVLKVKALMEFNPREVARNVFECDQVVKGEAGEVRLVHVKQSTDRDLRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRAVENETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFEKNCNDNI CODDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLEFALPFEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSQSWLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSQSWLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVMVVTSYKYNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 960
DB 917 QLELPVKYAVMVVTSYKYNFTASENTSRVMQHQYQVSNLQORSLSPLSLVFLVPV 976
QY 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDIP 1020
DB 977 RLNTQVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDIP 1036
QY 1021 FFGIQEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQET 1080
DB 1037 FFGIQEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAFVRSQET 1096
QY 1081 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1137
DB 1097 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1153

Tue Nov 9 12:56:12 2004

```

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 136

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQLKKSTLFLSMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQLKKSTLFLSMQYSEEF 196

181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256

241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 316

301 NFEALKTIONLREKIFAIBGTQTGSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAQ 360
DB 317 NFEALKTIONLREKIFAIBGTQTGSSSFEHEMSQEGFSAITNSGPLLSTVGSYDWAQ 376

361 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436

421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480
DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 496

481 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVYLF 540
DB 497 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVYLF 556

541 HGTSGSGISPSHSORIASKLSPLQVFGOSLSGGQDLTMDGLVDLTAVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSORIASKLSPLQVFGOSLSGGQDLTMDGLVDLTAVGAQGHVLLRSQ 616

601 PVLRVKAIMFEPNPREAVRNVECDQVVKGEAGEVRVCLHVQKSTRDLREGIOQSVVT 660
DB 617 PVLRVKAIMFEPNPREAVRNVECDQVVKGEAGEVRVCLHVQKSTRDLREGIOQSVVT 676

661 YDLALDSGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
DB 677 YDLALDSGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 736

721 SLVGTPLSAFAGNLRLPVLAEDAQRLEFALFPFKNCGNDNICQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFAGNLRLPVLAEDAQRLEFALFPFKNCGNDNICQDDLSITFSFMSLDCLVVG 796

781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSVRKYSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSVRKYSTLQNRQSRQSWRLACESASSTEV 856

841 SGALKSTSCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKLLKANVTSENMMPTNKTEF 900
DB 857 SGALKSTSCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKLLKANVTSENMMPTNKTEF 916

901 QLELPVKYAVYVMTVSHGVSTKYLNFASENTSRVMQHQVQVSNLQORSIPISLVFLVPV 960
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFASENTSRVMQHQVQVSNLQORSIPISLVFLVPV 976

961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036

1021 FFGQBEFNATLKGNSLDFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQGAFFVRSQTE 1080
DB 1037 FFGQBEFNATLKGNSLDFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQGAFFVRSQTE 1096

1081 KVEPFEPVNPPLIVGVSSVGGLLALLIITAALYKLGFFKQYKDMSEGGPPGAEPQ 1137
DB 1097 KVEPFEPVNPPLIVGVSSVGGLLALLIITAALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match 99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFLFGSNLRQOPQK 136

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQLKKSTLFLSMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKFVSTVMEQLKKSTLFLSMQYSEEF 196

181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 256
```

Db 197 RIHFTFKFQNNPNRSLVKPIITQLLORTHTATGIRKVVRELNITNGARKNAFKILWVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWOCDAVLVGEQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGGISPSHSQRIAGSKSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGGISPSHSQRIAGSKSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 900
Db 917 QLELPVKYAVYVVTSHGVSTKYNFTASNTSRVMQHOYQVSNLQORSPLISLVFLVPV 976
QY 961 RLNOTVMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1020
Db 977 RLNOTVMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1036
QY 1021 FFGIQQEENATLGNLSPDWYIKTSHNLLIYVSTABILNDSVFTLLPQGAFAVSQET 1080
Db 1037 FFGIQQEENATLGNLSPDWYIKTSHNLLIYVSTABILNDSVFTLLPQGAFAVSQET 1096
QY 1081 KVEPEFVNPLPLIVGSSVGGILLALITAAALYKLGFFKRYOKDMWSEGGPGABPQ 1137
Db 1097 KVEPEFVNPLPLIVGSSVGGILLALITAAALYKLGFFKRYOKDMWSEGGPGABPQ 1153

RESULT 6

US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069, 401C1
; CURRENT APPLICATION NUMBER: US/10/207,655

; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176

Query Match 99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCBPI 76
QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPOK 120
Db 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPOK 136
QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSBEF 180
Db 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSBEF 196
QY 181 RIHFTFKFQNNPNRSLVKPIITQLLORTHTATGIRKVVRELNITNGARKNAFKILWVI 240
Db 197 RIHFTFKFQNNPNRSLVKPIITQLLORTHTATGIRKVVRELNITNGARKNAFKILWVI 256
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNITIASKPPRDHVFQVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLISTVGSYDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAALIIILNRVQSLVGLGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQVSVCP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWOCDAVLVGEQOPWGRFGAALTIVLDVNGDKLTVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGGISPSHSQRIAGSKSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGGISPSHSQRIAGSKSPRLQYFGQSLSGGQDLTMDGLVDTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
Db 677 YDLALDSGRPHSAVFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 900
Db 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTET 916

QY	901	OLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVLPV	960
DB	917	QLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVLPV	976
QY	961	RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
DB	977	RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIQCDIP	1036
QY	1021	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQGGAFVRSQTET	1080
DB	1037	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQGGAFVRSQTET	1096
QY	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEQ	1137
DB	1097	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEQ	1153
RESULT 7			
US-09-902-481A-5			
; Sequence 5, Application US/09902481A			
; Publication No. US20030054440A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Shifman, Julia			
; APPLICANT: Mayo, Stephen			
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY			
; FILE REFERENCE: A-70586-1/RET/RMS/RMK			
; CURRENT APPLICATION NUMBER: US/09/902,481A			
; CURRENT FILING DATE: 2001-07-09			
; PRIOR APPLICATION NUMBER: US 60/216,600			
; PRIOR FILING DATE: 2000-07-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 5			
; LENGTH: 1137			
; TYPE: PRT			
; ORGANISM: Artificial sequence			
; FEATURE:			
; OTHER INFORMATION: synthetic			
US-09-902-481A-5			
Query Match 99.7%; Score 5855; DB 10; Length 1137;			
Best Local Similarity 99.1%; Pred. No. 0;			
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;			
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLYQCDYSTGSCPEI	60
DB	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLYQCDYSTGSCPEI	60
QY	61	RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFPLGSLNRQOPQ	120
DB	61	RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFPLGSLNRQOPQ	120
QY	121	FPEALRGCPEDSDIAPLDGSGSIIPDPRRKEFVSTVMEQLKSKTLFSLMOYSEEF	180
DB	121	FPEALRGCPEDSDIAPLDGSGSIIPDPRRKEFVSTVMEQLKSKTLFSLMOYSEEF	180
QY	181	RIHFTFKFQNNPNRSLVKPIITQLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	240
DB	181	RIHFTFKFQNNPNRSLVKPIITQLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	240
QY	241	TGDKFGDPLGYEDVIEADREGVIRVYVIGVDAPFRSEKSKQELNTASKPRPHVQFN	300
DB	241	TGDKFGDPLGYEDVIEADREGVIRVYVIGVDAPFRSEKSKQELNTASKPRPHVQFN	300
QY	301	NFEALKTQNLQREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGELLSTVGSYDWAG	360
DB	301	NFEALKTQNLQREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGELLSTVGSYDWAG	360
QY	361	GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAILLRNVQSLVLCAPRYQHIGLVAMFR	420
DB	361	GVFLYTSKEKSTFINMTVRDSDMDAYLGYAAAILLRNVQSLVLCAPRYQHIGLVAMFR	420

QY	421	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEOTRGQSVSVCPL	480
DB	421	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEOTRGQSVSVCPL	480
QY	481	PRGORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAAYILF	540
DB	481	PRGORARWOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAAYILF	540
QY	541	HGTSGSGISPSHSQISORLAGSKLSPRLQYFGQSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ	600
DB	541	HGTSGSGISPSHSQISORLAGSKLSPRLQYFGQSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ	600
QY	601	PVLURVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDRREGIQSVVT	660
DB	601	PVLURVKAIMEFNPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDRREGIQSVVT	660
QY	661	YDLALDSGRPHSAVFNETKSTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRINF	720
DB	661	YDLALDSGRPHSAVFNETKSTRQTVGLGTQTCETLKLQLPNCIEDPVSPIVLRINF	720
QY	721	SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCVGNNDICQDDLSITFSFMSLDCVLVG	780
DB	721	SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCVGNNDICQDDLSITFSFMSLDCVLVG	780
QY	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQORSWRLACESASSTEV	840
DB	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQORSWRLACESASSTEV	840
QY	841	SGALKSTSCSINHPIPFENSEVTFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
DB	841	SGALKSTSCSINHPIPFENSEVTFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
QY	901	QLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVLPV	960
DB	901	QLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLVLVLPV	960
QY	961	RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
DB	961	RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
QY	1021	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQGGAFVRSQTET	1080
DB	1021	FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQGGAFVRSQTET	1080
QY	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEQ	1137
DB	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEQ	1137
RESULT 8			
US-09-945-265-4			
; Sequence 4, Application US/09945265			
; Patent No. US20020123614A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy A.			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Lu, Chafen			
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A			
; FILE REFERENCE: CBN-002CP			
; CURRENT APPLICATION NUMBER: US/09/945,265			
; CURRENT FILING DATE: 2001-08-31			
; PRIOR APPLICATION NUMBER: US 60/229,700			
; PRIOR FILING DATE: 2000-09-01			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 4			
; LENGTH: 1152			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-945-265-4			

```
Query Match          99.6%; Score 5852.5; DB 9; Length 1152;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
QY 61 RLOQVPEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLCFLFGSNLRQPOK 120
Db 77 RLOQVPEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLCFLFGSNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKFEVSTVMEQLKSKTLFSLMOYSEBF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKFEVSTVMEQLKSKTLFSLMOYSEBF 196
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPDRDHVFOVN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPDRDHVFOVN 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCP 496
QY 481 PRGORARWQCDVILYCGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWQCDVILYCGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 555
QY 541 HGTSGSGISPSHQSRIAGSKSLSPRLQYFGOSLGGODLTMDGLVDLTVGAGHVLLRSQ 600
Db 556 HGTSGSGISPSHQSRIAGSKSLSPRLQYFGOSLGGODLTMDGLVDLTVGAGHVLLRSQ 615
QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVGKEAGEVRVCLHVOKSTRDLRREGIOISVVT 660
Db 616 PVLRVKAIEMFNPREVARNVFCNDQVVGKEAGEVRVCLHVOKSTRDLRREGIOISVVT 675
QY 661 YDLALDSGRPHRAVENETKSTRTOTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 720
Db 676 YDLALDSGRPHRAVENETKSTRTOTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLF 735
QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 780
Db 736 SLVGTPLSAFNLRPVLAEDAQLFTALPFPEKNCNDNICODDLSITFSFMSLCLVVG 795
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLRYKYSTIQLQORSQSWELACESASTEV 840
Db 796 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLRYKYSTIQLQORSQSWELACESASTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVTSHGVSSTKYNLFTASENTRVMQHQYQVSNLQORSLPISVFLVPV 960
Db 916 QLELPVKYAVVMVTSHGVSSTKYNLFTASENTRVMQHQYQVSNLQORSLPISVFLVPV 975
QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPHSDFLAELKAPVNVCSIAVCQRIQCDIP 1020
Db 976 RLNQTVIWRDPQVTFSENLSTCHTKERLPHSDFLAELKAPVNVCSIAVCQRIQCDIP 1035
QY 1021 FFGIQEENATLKGNSLFDWYIKTSHNLLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080
Db 1036 FFGIQEENATLKGNSLFDWYIKTSHNLLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1095
QY 1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFRQYKDMMSSEGPPGAEPQ 1137
Db 1096 KVEPFEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFRQYKDMMSSEGPPGAEPQ 1152

RESULT 9
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match          99.5%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
QY 61 RLOQVPEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLCFLFGSNLRQPOK 120
Db 61 RLOQVPEAVNMSLGLSLAATTPPQLLACGPTVHQTCSSENTYVKGCLCFLFGSNLRQPOK 120
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKFEVSTVMEQLKSKTLFSLMOYSEBF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKFEVSTVMEQLKSKTLFSLMOYSEBF 180
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
Db 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPDRDHVFOVN 300
Db 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPDRDHVFOVN 300
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
Db 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCP 480
QY 481 PRGORARWQCDVILYCGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 481 PRGORARWQCDVILYCGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
```

541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIOSVVT 660
601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIOSVVT 660
661 YDLALDSGRPHSAVNETKNSRTRQVGLTQTCTETLKLQLPNCIEDPVSPIVLRLNF 720
661 YDLALDSGRPHSAVNETKNSRTRQVGLTQTCTETLKLQLPNCIEDPVSPIVLRLNF 720
721 SLVGTPLSAFNGLRPVLAEADQRLFTALFPFKNCNGNDNICDDLSITFSFMSLDCLVVG 780
721 SLVGTPLSAFNGLRPVLAEADQRLFTALFPFKNCNGNDNICDDLSITFSFMSLDCLVVG 780
781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQNSORSWRWLACESASSTEV 840
781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQNSORSWRWLACESASSTEV 840
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900
901 QLELPVKYAVYVMVTVSHGVSTKYNFTASENTRVMQHOYQVSNLQORSLSISLVLVPV 960
901 QLELPVKYAVYVMVTVSHGVSTKYNFTASENTRVMQHOYQVSNLQORSLSISLVLVPV 960
961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020
961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTET 1080
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTET 1080
1081 KVEPFEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137
1081 KVEPFEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137

RESULT 10
US-09-902-481A-3
; Sequence 3, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3

Query Match 99.3%; Score 5832; DB 10; Length 1137;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGLSYQCDSYSGCEPI 60
Db 1 FNLDENAMTQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGLSYQCDSYSGCEPI 60

61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOK 120
61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSNTYVKGICFLFGSNLRQPOK 120
121 FPEALRGCEQSDIAFLIDGSGSIIPHDPRMKBEFVSTVMEOLKSKSKTLFSLMOYSEEF 180
121 FPEALRGCEQSDIAFLIDGSGSIIPHDPRMKBEFVSTVMEOLKSKSKTLFSLMOYSEEF 180
181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNINANGARKNAFKLILVI 240
181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELLNINANGARKNAFKLILVI 240
241 TDGEKFGDPLGYEDVPIEADREGVIRVIVGVGDAFSEKSEKSOELINTIASPPRDHVFQVN 300
241 TDGEKFGDPLGYEDVPIEADREGVIRVIVGVGDAFSEKSEKSOELINTIASPPRDHVFQVN 300
301 NFPAKXTIQLNOLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360
301 NFPAKXTIQLNOLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAWER 420
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAWER 420
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETOTRGGOVSVCP 480
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYETOTRGGOVSVCP 480
481 PRGQARWQCDVAVLYGSGQGWGFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
481 PRGQARWQCDVAVLYGSGQGWGFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSISGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600
601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIOSVVT 660
601 PVLRVKAIMFNPVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIOSVVT 660
661 YDLALDSGRPHSAVNETKNSRTRQVGLTQTCTETLKLQLPNCIEDPVSPIVLRLNF 720
661 YDLALDSGRPHSAVNETKNSRTRQVGLTQTCTETLKLQLPNCIEDPVSPIVLRLNF 720
721 SLVGTPLSAFNGLRPVLAEADQRLFTALFPFKNCNGNDNICDDLSITFSFMSLDCLVVG 780
721 SLVGTPLSAFNGLRPVLAEADQRLFTALFPFKNCNGNDNICDDLSITFSFMSLDCLVVG 780
781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQNSORSWRWLACESASSTEV 840
781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYRKVSTLQNSORSWRWLACESASSTEV 840
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900
901 QLELPVKYAVYVMVTVSHGVSTKYNFTASENTRVMQHOYQVSNLQORSLSISLVLVPV 960
901 QLELPVKYAVYVMVTVSHGVSTKYNFTASENTRVMQHOYQVSNLQORSLSISLVLVPV 960
961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020
961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTET 1080
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTET 1080
1081 KVEPFEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137
1081 KVEPFEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137

RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match 59.1%; Score 3473; DB 14; Length 1163;
Best Local Similarity 61.3%; Pred. No. 6e-284;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARFGOSVVOQLQGSRRVVGAPQBIIVANQSGSLYQCDYSGSCPEI 60
DB 20 FNLDTTELTAFRVDSAGFSDVSVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGFLCFLFGSNLRQOPQK 120
DB 80 GLQVPEAVNMSLGLSLASTTSQQLLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137
QY 121 FPAALRGCTQEDSDIAFLDGGSSIIIPDFRRMKEFVSTVMEQLKKSKTLFSLMOWSEEF 180
DB 138 LPVSROECRQEQDIFVLDDGSSISRRNFATMNFVRAVISQFQPSQFSLMQFSNKF 197
QY 181 RIHFTPEFQNNPNSLAKPTIOLGRTHATGVRKVIRELNIINGARKNAFKILIVI 240
DB 198 QTHFTPEEFRTSNPLUSLASVHLQGLQFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257
QY 241 TDCEKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVFOYN 300
DB 258 TDGKKGDSLDYKVTIPMADAGIIRYALGVGLAFQNRNSWKNELNDIAKPSQEHIFKVE 317
QY 301 NFEALKTIONQREKIFAIEGTQSSSFHEMSQEGFSAITSNGLPLSTVGSYDWAQ 360
DB 318 DFDALDKIQNLKEKIFAIEGTETSSSSFELEMAQEGFSAVTPDGPVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAIILNRNVQSLVLCAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINNSQBNVMDRDSYGLSTELALWKGVQSLVLCAPRYQHTKAVIPT 437
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSDVDVDSNGSTDLVLICAPHYEQTRGGQVSVCP 480
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSDVDVDTGSDTLVLICAPHYEQTRGGQVSVCP 497
QY 481 PRGQARQOCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRCVAVLF 540
DB 498 PRGWR-RWMCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVVLGAPGEENRCAVILP 556
QY 541 HGTSGSGSLPSHSORIASKSLSPLOYGOSLSGGQDLTMGLVDLTVGAQHVLRLRSQ 600
DB 557 HGVLGPSLPSHSORIASKSLSSRLQYFGQALSGGQDLTDGLVDLAVGARGQVLLRTR 616
QY 601 PVLRYKAIMFEENPREVARNEFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 617 PVLWVGUSMQFIPAEIPSAFEQREQVVSQTLVQSNICLIDKRSKNLLGSRDLQSSVT 676

QY 661 YDLALDSGRPHSRVAFVNETKSTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVILRLNF 720
DB 677 LDALDPGRLSPRATFQETKNRSLSRVVLGLKAHCENFNLLPSCVEDSVTPITLRLNF 736
QY 721 SLVGTPLSAFGNLAPVLAEDARLFTALFPEKCKGNDNI CODDLSITSFMSLDCLVVG 780
DB 737 TLVGKPLLAFLNLRPLMLAALAQRYFTASLPPEKCKGADHICQDNLGISFSFPLKSLLVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYKVKVSTLQNRQORSWRLACESASSTEV 840
DB 797 SNLEINAEVWVNDGEDSYGTTITFSPAGLSYRYVABGQKQGRSLRSLHLTCD SAPVG-- 854
QY 841 SGALKSTSCSINHPIFENSEVETNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
DB 855 SQGTWSTSCRINHILIFRGAQITELATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFTAS-ENTSRVMOHQYOVNGLGORSILPISLVFLVP 959
DB 915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974
QY 960 VRLNQTVINDRPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVCSIAVCORIOCDI 1019
DB 975 VELNQEAVMWVDEVSHPQNPESLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRPRCDV 1034
QY 1020 PFTGQIEFENATLKNLSFDWYIKTSHNHLIIVSTABILEFENDSVFTLLPGGAFVRQTE 1079
DB 1035 PSFQVQBELDTLKNLSFGWVQIILQKVVSVVSAEITFTSVYSQLPQGEAFMRAQTT 1094
QY 1080 TKVEPEVFNPLIYVSSVGGILLALITAAIYKLGFFKRYQKDMME 1128
DB 1095 TVLEKYKVNPTPLIYVSSIGGLLLALITAVLYKVGFFKRYQKEMEE 1143

RESULT 12

US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4

Query Match 58.7%; Score 3450; DB 9; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.3e-282;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARFGOSVVOQLQGSRRVVGAPQBIIVANQSGSLYQCDYSGSCPEI 60
DB 20 FNLDTTELTAFRVDSAGFSDVSVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGFLCFLFGSNLRQOPQK 120
DB 80 GLQVPEAVNMSLGLSLASTTSPPQLLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137

Publication No. US20030077278A1
 GENERAL INFORMATION:
 APPLICANT: Michael W.
 APPLICANT: Van der Vieren, Monica
 TITLE OF INVENTION: No. US20030077278A1 Human 2
 FILE REFERENCE: 27866/35004
 CURRENT APPLICATION NUMBER: US/09/891,943
 CURRENT FILING DATE: 2001-06-26
 PRIOR APPLICATION NUMBER: 09/193,043
 PRIOR FILING DATE: 1998-11-16
 PRIOR APPLICATION NUMBER: 08/286,889
 PRIOR FILING DATE: 1994-08-05
 PRIOR APPLICATION NUMBER: 08/362,652
 PRIOR FILING DATE: 1994-12-21
 PRIOR APPLICATION NUMBER: 08/943,363
 PRIOR FILING DATE: 1997-10-03
 NUMBER OF SEQ ID NOS: 114
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 4
 LENGTH: 1163
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-891-943-4

Query Match 58.7%; Score 3450; DB 10; Length 1163;
 Best Local Similarity 61.0%; Pred. No. 5.3e-282;
 Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

Qy 1 FNDTENAMTFQENARGFGOSVVLQGSVVGAPQIIVAAQNRGSLYQCDYSTGSCPEI 60
 Db 20 FNDTEELTAFRVDSAGFGSVVQYANSVVGAPQIIVAAQNRGSLYQCDYSTGSCPEI 79
 Qy 61 RLQVPEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLPFLGSLNRQPOK 120
 Db 80 GLQVPEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLPFLGSLNRQPOK 137
 Qy 121 FPEALRCQPDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQKSKTSLMOYSEEF 180
 Db 138 LPVSRQCPREQDQIVFLIDGSGSISRRFATMNFVRAVISQRPSTQSLMOYSEEF 197
 Qy 181 RIHTEFEFQNNPRLSVKPIITQLGRTHPTATGVKRVIRELLNITNGARKNAFKILIVI 240
 Db 198 QTHTEFEFRTSNPLSLASVHQLQGYTATATQNVVHRLFHASYGARRDAIKILIVI 257
 Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSEQLNTIASKPRDHVQVN 300
 Db 258 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSEQLNTIASKPRDHVQVN 317
 Qy 301 NFEALKTQNLREKIPAIETGTGSSSSPEHEMSQEGSAITNSGPLLSTVGSYDWAG 360
 Db 318 DFDALKDQNLREKIPAIETGTGSSSSPEHEMSQEGSAITNSGPLLSTVGSYDWAG 377
 Qy 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 Db 378 GAFLYPPNMSPTFNNSENVMDRDSYLGYSTELALWKGVQSLVGLGAPRYQHIGLVAMFR 437
 Qy 421 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQSVCP 480
 Db 438 QVSRQWRMKAIEVIGTQISYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQSVCP 497
 Qy 481 PRGQARWQCDVLYGEOQPMRFGAALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 540
 Db 498 PRGWR-RWMCDAVLYGEOQPMRFGAALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 556
 Qy 541 HGTSGSGISPSHSQRIAGSKISPRIQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 Db 557 HGVLPSPISPSHSQRIAGSKISPRIQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
 Qy 601 PVLRYKAIMFNPREVARNVFECDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
 Db 617 PVLWVGVSQMFIPEIPRSAFEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
 Qy 661 YDLALDSGRPHSAFVNETKNSTRRTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720

Qy 121 FPEALRCQPDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQKSKTSLMOYSEEF 180
 Db 138 LPVSRQCPREQDQIVFLIDGSGSISRRFATMNFVRAVISQRPSTQSLMOYSEEF 197
 Qy 181 RIHTEFEFQNNPRLSVKPIITQLGRTHPTATGVKRVIRELLNITNGARKNAFKILIVI 240
 Db 198 QTHTEFEFRTSNPLSLASVHQLQGYTATATQNVVHRLFHASYGARRDAIKILIVI 257
 Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSEQLNTIASKPRDHVQVN 300
 Db 258 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSEQLNTIASKPRDHVQVN 317
 Qy 301 NFEALKTQNLREKIPAIETGTGSSSSPEHEMSQEGSAITNSGPLLSTVGSYDWAG 360
 Db 318 DFDALKDQNLREKIPAIETGTGSSSSPEHEMSQEGSAITNSGPLLSTVGSYDWAG 377
 Qy 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
 Db 378 GAFLYPPNMSPTFNNSENVMDRDSYLGYSTELALWKGVQSLVGLGAPRYQHIGLVAMFR 437
 Qy 421 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQSVCP 480
 Db 438 QVSRQWRMKAIEVIGTQISYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQSVCP 497
 Qy 481 PRGQARWQCDVLYGEOQPMRFGAALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 540
 Db 498 PRGWR-RWMCDAVLYGEOQPMRFGAALTVLGDVNGDKLTDVAIGAPDEENRGAVYLF 556
 Qy 541 HGTSGSGISPSHSQRIAGSKISPRIQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
 Db 557 HGVLPSPISPSHSQRIAGSKISPRIQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
 Qy 601 PVLRYKAIMFNPREVARNVFECDQVVGKEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660
 Db 617 PVLWVGVSQMFIPEIPRSAFEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
 Qy 661 YDLALDSGRPHSAFVNETKNSTRRTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
 Db 677 LDALAFGRSPRAIFQETKRSLSRVVGLGKACENFNLLPSCVEDSVIPILRLNF 736
 Qy 721 SLVCTPLSAFNLPLAEDAQRLFTALPPEKNCNDNIQDDLSITFFSFMSLDCLVVG 780
 Db 737 TLVGKPLAFLNPLAALAQRIFTASLPPEKNCADHIQDNLGIFSFPGLKSLIVG 796
 Qy 781 GPRFNVTVVRNDEGSYRTQVTFPFLDLSYKSVSTLQNRORSWRLACESASSTEV 840
 Db 797 SNLELNAEVMVWNGEDSYGTTITFHPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854
 Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
 Db 855 SGTWSTSCRINHLIFRGAQITFATEDVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914
 Qy 901 QLELPVKYAVMVVTSVSTKYLNFAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 959
 Db 915 QLELPVKYAVMVVTSVSTKYLNFAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 974
 Qy 960 VELNQTIVWRDQVTFSENLSTCHTKERLPKSHDFLAEKAPVNVGSIACVQRIQCDI 1019
 Db 975 VELNQEAVMVDEVSHPPQNPRLSRLCSSEKTIAPASDFLAHIQNPVLDLCSIAGLFRCDV 1034
 Qy 1020 PFGIOEBFNATLKNLSPDWIKTSHNHLITVSTAEIIPNDVSFTLLPQGAFFRSOTE 1079
 Db 1035 PSFSVQEEJDLFTLKNLSEFGWRQILQKVSVVSVVAEIIIDTSVYSQLPQGAFFRAQTI 1094
 Qy 1080 TKVEPEVNPPLIIVGSSVGLLILALITAAALYKLGFFKRYQKDMWSE 1128
 Db 1095 TVLEKYKVNPIPLIVGSSIGGLLILALITAVLYKVGFFKRYQKEMWEE 1143

Db 677 LDALAPGLSPRAIFQETKRSLSRVRLGLKAHCENFNLLPSCVEDSVIPIILRLNF 736
Qy 721 SLVGTPLSAFNGRLVLAEDAQRLLTALFPPEKNCNDNI CODDLSITSPMSLDCLVVG 780
Db 737 TLVGKPLAFNRLPMLAALQRYETASLPPEKNCADHI CODNIGISFSFPLKSLVVG 796
Qy 781 GPRFNVTVVRNDEGDSYRTQVTPFFPLDLSYRKVSTLQORSORSWELACESASSTEV 840
Db 797 SNLEINAEVWMDGSDSYGTTIIFSHPAGLSYRVVAEGQKQQLRSLHLTC--CSAPVG 854
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 855 SQGTWSTSCRINHPIFRGAQITFLATPDVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914
Qy 901 QLELPVKYAVVMVTSYHGVSTKYLNFTAS-ENTSRVMOHQVOVSNLQORSILPISLVFLVP 959
Db 915 QLELPVKYAVVIVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSEINFWP 974
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVCSIAVCORICDI 1019
Db 975 VELNQEAVMDVESHQNPNSLRCSSEKIAPPASDFLAHQKNPVLDSCSIAGCURFRCDV 1034
Qy 1020 PFGIOEFNATLKNLSFDWIKTSNHLIIVSTAEILFNDVSFTLLPGCGAFVRSQTE 1079
Db 1035 PSFSVQBELDFTLKNLSFGWVRQILQKYSVSVSAEIIIFDTSVYSQLPQGEAFMRAQTI 1094
Qy 1080 TKVEPFEVNPPLPIVGVSSVGGILLLALITAAALYKLGFFKROKDMME 1128
Db 1095 TVLEKYKVNPIPLIVGSSIGGLLLALITAVLYKGVFFKROKYMEE 1143

RESULT 14

US-09-350-259-2
; Sequence 2, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-2

Query Match 58.2%; Score 3417; DB 9; Length 1161;
Best Local Similarity 59.8%; Pred. No. 3.3e-279;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;
Qy 1 FNLDTENAMTFQENARGGOSVQVQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPTI 60
Db 17 FNLVDDEPTIFQEDAGGFGQSVVQGGSRVVVGAPLEVVAAHQSLYQCDYSTGSCPTI 76
Qy 61 RLQVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVVKGLCFPLFGNSRQOQPK 120
Db 77 PLHTRPEAVNMSLGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-EIITQ 135

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKEFVSTVMEQLKSKTFLSLMOYSBEF 180
Db 136 VPDATPECPHQEMDVLFLIDGSGSIDQNDQMGKGFVQAVMQCFEGTDTLTFALMOYSNLL 195
Qy 181 RIHTEFEFONNPRSLAVKPIITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
Db 196 KIHFTFTQFRTPSQQLVDPVQLKGLTFTATGLTAVTQLFHHKNGARKSAKKILIVI 255
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVIGVDAPRSEKSRQELNTIASKPRDHVFOVN 300
Db 256 TDGQKYKDPLEYSVVIQAEKAGIIRVAIGVGHAFQGPTRQELNTISSAPQDHVFKVD 315
Qy 301 NFEALKITONLRKPIFAIEGTQGTSSSPSEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db 316 NFAALGSIQLOQEKIYAVEGTQSRASSSQFHEMSQEGFSTALTMDGLFLGAVGSFWSG 375
Qy 361 GVRLYTSKESKSTFNMTFRVSDMDNDAYLGVAALILNRVQSLVIGAPRYOHIGLAVMFR 420
Db 376 GAFLYPPNMGPTFNNMQENVMDRDSYLGYSTELALWGVQNLVIGAPRYOHTKAVIFT 435
Qy 421 QNTGMWESNANVKGTQIYGAYFGASLCSVDVDSNGSTDLVIGAPHYVECTRGGOVSCPL 480
Db 436 QVSRQWRKKAETVCTQIGSYFGASLCSVDVDSNGSTDLVIGAPHYVECTRGGOVSCPL 495
Qy 481 PRQBARWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 496 PRQVRVQWQCDVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
Qy 541 HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAGHVLRLRSQ 600
Db 556 HGASEGSISSHSORLAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAGHVLRLRSQ 615
Qy 601 PVLKVAIMEFNPREVARNVFECDQVVGKZAGEVRCVCLHVOKSTRDRREGIOISVVT 660
Db 616 PVLKVGVMRFSPEVAKAVYRCWEKPSALEAGDATVCLTIQKSSDQL--GDIOSSVR 673
Qy 661 YDLALDSGRPHSAVNEFNKSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVILRLNF 720
Db 674 FDLALDPGLTSAIRAFNETKNPTLRRKTLGLGHCEKTLKLLPDCVEDVVSIIHLNF 733
Qy 721 SLVGTPLSAFNGRLVLAEDAQRLLTALFPPEKNCNDNI CODDLSITSPMSLDCLVVG 780
Db 734 SLVREPIPSQNLRLPVLAVGSDQLFTASLPFFKNCQDGLCEGDLGVTLSFSLQTLTVG 793
Qy 781 GPRFNVTVVRNDEGDSYRTQVTPFFPLDLSYRKVSTLQORSORSWELACESASSTEV 840
Db 794 SSLELNVIVTVNAGEDSYGTVVSLYYPAGLSHRRVSGAQKPHQSAURLACETV-PTED 852
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGKLLKANVTSENMPRTNKTEF 900
Db 853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTPDVSVYKATLGDRLMRASSSENKASSKATF 911
Qy 901 QLELPVKYAVVMVTSYHGVSTKYLNFTAS-ENTSRVMOHQVOVSNLQORSILPISLVFLVP 959
Db 912 QLELPVKYAVVIVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSEINFWP 971
Qy 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHDFLAELRKAPVNVCSIAVCORICDI 1019
Db 972 VLLNGVAVMDVMEAPSQSL--PCVSEKRPQHSDFLTQISRPMLDSCIADCLQFRCDV 1029
Qy 1020 PFGIOEFNATLKNLSFDWIKTSNHLIIVSTAEILFNDVSFTLLPGCGAFVRSQTE 1079
Db 1030 PSFSVQBELDFTLKNLSFGWVRQILQKVLVWSVAEITFDTSVYSQLPQGEAFMRAQME 1089
Qy 1080 TKVEPFEVNPPLPIVGVSSVGGILLLALITAAALYKLGFFKROKDMME 1128
Db 1090 MVLREDEVNAIPIIMSGSSVGALLLALITATLYKLGFFKRYKEMLED 1138

RESULT 15

US-09-891-943-2
; Sequence 2, Application US/09891943
; Publication No. US2003007728A1

Tue Nov 9 12:56:12 2004

GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-2

Query Match 58.2%; Score 3417; DB 10; Length 1161;
Best Local Similarity 59.8%; Pred. No. 3.3e-279;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

QY 1 FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLVVEPTTFQEDAGGQSVVQFGGSRVVGAPLEVAANQGRLYDCAAAATGMCQPI 76

QY 61 RLQVPVAVNMGLSLAATSPQQLACGTVHTQTSENTYVVKGLCLFLFGLNLRQPOK 120
Db 77 PLHTRPAVNMGLTAASTNGSRLLACGPTLHRVCGENSYSKSGCLLLGSRW-EIIQT 135

QY 121 PPEALRGCPQEDSDIARLIDGSGTIIPDRFRMKEFVTVMEQKKSKTLFSLMOYSEEP 180
Db 136 VPDATPECPQEMDVLFLIDGSGSDQDNFQMGKFGVQAVMGQPEGDTTLFALMOYSNLL 195

QY 181 RIHFTFKFONNPRSLVKPITQLLGRHTHTATGVRKVRIRRELLNTNGARKNAKILIVI 240
Db 196 KIHFTFTQRTSPSQSLVDPIVLQKGLTFTATGTLTVVITQLFHHKNGARKSAKKILIVI 255

QY 241 TDGEKFGDPLGHEDVVPADREGVIRYVIGDAPRSEKSOELNTIASKPRDHVFOVN 300
Db 256 TDGQYKDPLEYSYDIPQAEKAGIIRYAGVGHAFQGTARQELNTIISAPPQDHVFKVD 315

QY 301 NFEALKTIQNLREKIFALEGTOTGSSSFHEMSQEGFSAAITSNGPLLLSTVGSYDWAG 360
Db 316 NPAALGSIQKQKQKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGFSWSG 375

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVQSLVLGAPRYQHIGLVAMPR 420
Db 376 GAFLYPPNMPPTFINMSQENVDMRDSYLSYSTELALWKGVQNLVLGAPRYQHTGKAVPT 435

QY 421 QNTGWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLICAPHYEOTRGQSVSCPL 480
Db 436 QVSROWRKKAETVGTIGSYFGASLCSVDVDSNGSTDLLIGAPHYEQTRGQSVSCPL 495

QY 481 PRGQARWOCDAVLYGEOQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 496 PRGQRVQWOCDAVLRGEQHPWGRFGAALTVLGDVNEKDLIDVAIGAPGEQNRGAVYLF 555

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMGDLVLTGQAQGHVLLLRQ 600
Db 556 HGASESGISPSHSQRIASSQLSPRLQYFGQALSGGQDLTQDGLMDLAVGARGQVLLLRSL 615

QY 601 PVLRVKAINMEFPNPREVARNVFECDQVVKGEAGEVVCILHYOKSTDRRLREGQIOSVVT 660
Db 616 PVLKVGAMRFSPEVEKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

QY 661 YDLALDSGRPHSRVAFNETKNSRRQTVLGLTQTCETLKLQLPNCIEDPVSFVILRLNF 720

Search completed: November 9, 2004, 12:47:03
Job time : 102.5 secs

Db 674 FDLALDPGLTSRAIPNETKNPTLTRKTLGLGHCETLKLPLPCDVEDVSPILLHLNF 733

QY 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFFKNCGNDNICQDDLSTFMSLDCLVVG 780

Db 734 SLVREFIPSPQNLRFVLAVGSQDLFTASLPFFKNCGQDGLCEGLGVTLSFSGLTITVG 793

QY 781 GPREFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRORSQSWRLACSSASSTEV 840

Db 794 SSLELNVITVMNAGSDSYGTVVSLLYPAGLSHRRVSGAQKQPHQSALRIACETV-PTD 852

QY 841 SGALKSTCSINHPPIFENSEVTNITEDVDSKASLGNKLLKANKVNTSENMMPTNKTEF 900

Db 853 EG-LRSSRCVNHPIFHEGSGNGTFIVTFDVSYKATLGDRLMRASASSENKASSKATF 911

QY 901 QLELPVKYAVYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQQRSLPISLFLVP 959

Db 912 QLELPVKYAVYTMISRQESTKYFNATSDKKMKEAHRVYRVNLSORDLAISINFWVP 971

QY 960 VRLNQTVINDRPOQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIOCDI 1019

Db 972 VLLNGVAVDVMVMEAPSQL--PCVSEKPPQSHSDFLTQISRSPMLDCSIADCLQFRCDV 1029

QY 1020 PFFGIQBEFNATLKGNLSPDWYIKTSHNHLIIVSTAELFENDSVFTLLPQGGAFVRSQTE 1079

Db 1030 PSFSVQEEELDTLKGNLSPGWVRETLOKKVLVSVVAEITDTSVYSQLPQGEAFMRAQME 1089

QY 1080 TKVEPPFVNPPLPIVGVSSVGGLLLALITAAALYKLGFFKRYQKDMSE 1128

Db 1090 MVLEEDVYNAIPIIMGSSVGAALLLALITATLYKLGFFKRYKHEMLED 1138

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds
(without alignments)
3278.416 Million cell updates/sec

Title: US-09-902-481B-6
Perfect score: 5875
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMGEGGPGGAEPQ 1137

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5868	99.9	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5868	99.9	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5868	99.9	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5868	99.9	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5868	99.9	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5868	99.9	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5868	99.9	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5868	99.9	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5868	99.9	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5837.5	99.4	1152	2 US-08-476-062A-43	Sequence 3, Appli
12	5837.5	99.4	1152	5 PCT-U596-01314-43	Sequence 43, Appl
13	5837.5	99.4	1152	6 5424399-2	Sequence 43, Appl
14	3473	59.1	1163	2 US-08-476-062A-44	Sequence 44, Appl
15	3473	59.1	1163	5 PCT-U596-01314-44	Sequence 44, Appl
16	3450	58.7	1163	1 US-08-173-497-4	Sequence 4, Appli
17	3450	58.7	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3450	58.7	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3450	58.7	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3450	58.7	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3450	58.7	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3450	58.7	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3450	58.7	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3450	58.7	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3450	58.7	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3417	58.2	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3417	58.2	1161	1 US-08-286-889-2	Sequence 2, Appli

28	3417	58.2	1161	1 US-08-485-618-2	Sequence 2, Appli
29	3417	58.2	1161	1 US-08-362-652-2	Sequence 2, Appli
30	3417	58.2	1161	2 US-08-605-672-2	Sequence 2, Appli
31	3417	58.2	1161	2 US-08-482-293A-2	Sequence 2, Appli
32	3417	58.2	1161	2 US-08-943-363-2	Sequence 2, Appli
33	3417	58.2	1161	3 US-09-193-043-2	Sequence 2, Appli
34	3417	58.2	1161	4 US-09-688-307A-2	Sequence 2, Appli
35	3417	58.2	1161	4 US-09-350-259-2	Sequence 2, Appli
36	3401.5	57.9	1161	1 US-08-485-618-99	Sequence 99, Appl
37	3401.5	57.9	1161	2 US-08-605-672-99	Sequence 99, Appl
38	3401.5	57.9	1161	2 US-08-482-293A-99	Sequence 99, Appl
39	3401.5	57.9	1161	2 US-08-943-363-99	Sequence 99, Appl
40	3401.5	57.9	1161	3 US-09-193-043-99	Sequence 99, Appl
41	3401.5	57.9	1161	4 US-09-688-307A-99	Sequence 99, Appl
42	3401.5	57.9	1161	4 US-09-350-259-99	Sequence 99, Appl
43	3239.5	55.1	1161	3 US-09-193-043-55	Sequence 55, Appl
44	3239.5	55.1	1161	4 US-09-688-307A-55	Sequence 55, Appl
45	3239.5	55.1	1161	4 US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1
US-08-173-497-3
; Sequence 3, Application US/08173497
; Patent No. 5437958
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van Der Vieren, Monica
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha
; TITLE OF INVENTION: Subunit
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 S. Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/173,497
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5437958and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-173-497-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGSVVQLQSRVVVGAPQBIIVANQSGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARGFGSVVQLQSRVVVGAPQBIIVANQSGSLYQCDYSTGSCEPI 76

RESULT 2
US-08-286-889-3
Sequence 3, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: P38,659
REFERENCE/DOCKET NUMBER: 27865/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-889-3
Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTVYVKGCLFLFGSNLRQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTVYVKGCLFLFGSNLRQPOK 136
QY 121 FPEARLGCQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKSKTLFSLMOYSEEF 180
Db 137 FPEARLGCQEDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEOLKSKTLFSLMOYSEEF 196
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELNTINGARKNAFKILVI 240
Db 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNTINGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 316
QY 301 NFPAALKTIONQLEKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360
Db 317 NFPAALKTIONQLEKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 376
QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQVSYCPL 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQVSYCPL 496
QY 481 PRGORARWQCDVLYGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 540
Db 497 PRGORARWQCDVLYGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 556
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGAGCHVLLRSQ 600
Db 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGAGCHVLLRSQ 616
QY 601 PVLRVKAIMFPNPREVARNFECDQVVKYKAGEVRVCLHVQKSTRDRREGQIQSVVT 660
Db 617 PVLRVKAIMFPNPREVARNFECDQVVKYKAGEVRVCLHVQKSTRDRREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETHKSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVRLNF 720
Db 677 YDLALDSGRPHSRAVFNETHKSTRQTOVLGTOTCETLKLQLPNCIEDPVSPIVRLNF 736
QY 721 SLVGTPLSAFGNLRPVLAEDAQRLEFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEDAQRLEFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVRNDGEDSVRTQVTFPFDLDSYRKVSTLONORSORSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSVRTQVTFPFDLDSYRKVSTLONORSORSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITPDVDSKASLGNKLLKANVTSENMPNINKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTENITPDVDSKASLGNKLLKANVTSENMPNINKTEF 916
QY 901 QLELPVKAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 960
Db 917 QLELPVKAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 976
QY 961 RLNQTVIMDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020
Db 977 RLNQTVIMDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1036
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAILFNDVSFTLLPGQAFVRSQTET 1080
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAILFNDVSFTLLPGQAFVRSQTET 1096
QY 1081 KVEPFFVNPPLPIVSSVGGLLLLALITAAALYKLGFFFKQYKDMMSGGPPGABPQ 1137
Db 1097 KVEPFFVNPPLPIVSSVGGLLLLALITAAALYKLGFFFKQYKDMMSGGPPGABPQ 1153

361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
421 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVQVCP 480
437 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVQVCP 496
481 PRGORARWQCDVLYGEGQDPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVYLF 540
497 PRGORARWQCDVLYGEGQDPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVYLF 556
541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 600
557 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 616
601 PVLRVKAIMENPREVARNPFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
617 PVLRVKAIMENPREVARNPFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
661 YDLALDSGRPHSRVAFNETKNSTRQVGLTQTCETLKLQPCNIEDPVPVILRLNF 720
677 YDLALDSGRPHSRVAFNETKNSTRQVGLTQTCETLKLQPCNIEDPVPVILRLNF 736
721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 780
737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRSWLACESSASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRSWLACESSASSTEV 856
841 SGALKSTSCSINPIPIENSEVTFNITFDVDSKASIGNKLLKLLKANTVSENMPRTKTEF 900
857 SGALKSTSCSINPIPIENSEVTFNITFDVDSKASIGNKLLKLLKANTVSENMPRTKTEF 916
901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHOYOVSNLQORSIPISLVFLVPV 960
917 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHOYOVSNLQORSIPISLVFLVPV 976
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQCIP 1020
977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCORIQCIP 1036
1021 FFGIOEEFNATLKNLSFDWYIKTSHNHLIYSTABILFNDVSFTLLPGQAFVRSQTET 1080
1037 FFGIOEEFNATLKNLSFDWYIKTSHNHLIYSTABILFNDVSFTLLPGQAFVRSQTET 1096
1081 KVEPEFVNPPLVLGVSSVGGLLALLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1137
1097 KVEPEFVNPPLVLGVSSVGGLLALLALITAAALYKLGFFKRYQKDMSEGGPPGAEPQ 1153

RESULT 3

US-08-485-618-3
; Sequence 3, Application US/08485618
; Patent No. 5728533
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Marsh, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; City: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,618
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32797
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-485-618-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQBNARFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQBNARFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQCSTYVYKGLCFI FGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHQCSTYVYKGLCFI FGSNLRQOPQK 136
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 180
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 196
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVI IRELLNITNGARKNAFKILVI 240
Db 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVI IRELLNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSKSRQEIANTIASKPPRHVFQV 300
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSKSRQEIANTIASKPPRHVFQV 316
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436
QY 421 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVQVCP 480
Db 437 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYEQTREGGVQVCP 496
QY 481 PRGORARWQCDVLYGEGQDPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGORARWQCDVLYGEGQDPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 600

REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-362-652-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;
Best Local Similarity 99.8%; Pred. No 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVLOGSRVVVVGAPQEIIVAAQORGSILYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVLOGSRVVVVGAPQEIIVAAQORGSILYQCDYSTGSCPEI 76
QY 61 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOQOK 120
DB 77 RLQVPVAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOQOK 136
QY 121 FPEALRCPOEDSDIAFLIDSGSIIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 180
DB 137 FPEALRCPOEDSDIAFLIDSGSIIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKEFQNNPNRSLVKPIITQLLGRTHATGVRKIVRELNITNGARKNAFKILIVI 240
DB 197 RIHFTFKEFQNNPNRSLVKPIITQLLGRTHATGVRKIVRELNITNGARKNAFKILIVI 256
QY 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKSOELNTIASKPPRHHVFOVN 300
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKSOELNTIASKPPRHHVFOVN 316
QY 301 NFEALKTONLREKIFAIEGTQTSSSSPHEMSOEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTONLREKIFAIEGTQTSSSSPHEMSOEGFSAAITNSGPLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQVSVCLP 480
DB 437 QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDLVLICAPHYYEOTRGQVSVCLP 496
QY 481 PRGQARWQCDVAVLYGEOQPGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVAVLYGEOQPGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPLRYQFQCSLGSQDQDLTMQGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRIAGSKLSPLRYQFQCSLGSQDQDLTMQGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMENPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNENKSTRQTOVLGTQTCETLKLQPLNCIEDPVSPVILRLNF 720
DB 677 YDLALDSGRPHSRAVFNENKSTRQTOVLGTQTCETLKLQPLNCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 796
QY 781 GPREFNVTVVNDGEDSYRTQVTFPPFLDLSVRKYSTLQONORSORSWELACESASSTEV 840

557 HGTSGSGISPSHSQRIAGSKLSPLRYQFQCSLGSQDQDLTMQGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
DB 617 PVLRVKAIMENPREVARNVFECNDQVVKKEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNENKSTRQTOVLGTQTCETLKLQPLNCIEDPVSPVILRLNF 720
DB 677 YDLALDSGRPHSRAVFNENKSTRQTOVLGTQTCETLKLQPLNCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 780
DB 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 796
QY 781 GPREFNVTVVNDGEDSYRTQVTFPPFLDLSVRKYSTLQONORSORSWELACESASSTEV 840
DB 797 GPREFNVTVVNDGEDSYRTQVTFPPFLDLSVRKYSTLQONORSORSWELACESASSTEV 856
QY 841 SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKXANTSENMMPTKNTKEF 900
DB 857 SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKXANTSENMMPTKNTKEF 916
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 960
DB 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 976
QY 961 RLNGTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQIQDIP 1020
DB 977 RLNGTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQIQDIP 1036
QY 1021 PFGIOEFNATLKNLSPDWIKTSHHLLIVSTAEILFNDVSFTLLPGQGAFAVRSOTET 1080
DB 1037 PFGIOEFNATLKNLSPDWIKTSHHLLIVSTAEILFNDVSFTLLPGQGAFAVRSOTET 1096
QY 1081 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGPFKROYKDMSEGGPPGAEPO 1137
DB 1097 KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGPFKROYKDMSEGGPPGAEPO 1153

RESULT 4
US-08-362-652-3
Sequence 3, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.

Db 797 GPEFNVTVVRNDEGDSYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLFLVFPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLFLVFPV 976
QY 961 RLNOTVIWDRPOVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1020
Db 977 RLNOTVIWDRPOVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1036
QY 1021 PFGIQEENATLKGNSLFDWIKTSHNHLIIIVSTAELFNDSVFTLLPGQGAFFVSOTET 1080
Db 1037 PFGIQEENATLKGNSLFDWIKTSHNHLIIIVSTAELFNDSVFTLLPGQGAFFVSOTET 1096
QY 1081 KVEPFEVNPPLIVGSSVGGILLIALLITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLIVGSSVGGILLIALLITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3
; Sequence 3, Application US/08605672
; Patent No. 5817515
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,672
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 ENLDTENAMTQENARGGQSVVQOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCERI 60
Db 17 ENLDTENAMTQENARGGQSVVQOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGSCERI 76
QY 61 RLQVPEAVNMSLGLSLAATSPPOLLIACGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATSPPOLLIACGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTSLTSMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTSLTSMQYSEEF 196
QY 181 RIHFTFKFQONNPRSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQONNPRSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSKOEELNTIASKPPRDHVFQV 300
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAFRSEKSKOEELNTIASKPPRDHVFQV 316
QY 301 NFEALKTIQNLREKIFAIEGTQGTSSSSPEHEMSQEGFSAATITNSGPLISTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQGTSSSSPEHEMSQEGFSAATITNSGPLISTVGSYDWAG 376
QY 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGVAAILLENRVQSLVGLAPRYQHIGLVAMER 420
Db 377 GVFLYTSKESKSTFNMTRVDSMDNDAYLGVAAILLENRVQSLVGLAPRYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 496
QY 481 PRQQRARWQCDVLYGEGQGPWGRFGAALTVDLGVNKGDKLTDVAIGAPCEEDNRGAVYLF 540
Db 497 PRQQRARWQCDVLYGEGQGPWGRFGAALTVDLGVNKGDKLTDVAIGAPCEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDLTVAQGHVLLRSQ 616
QY 601 PVLRVKAIIMEFNPREVARNVFCNDQVVKGEAGEVVRVCLHVQKSTRDLREGIQSVVT 660
Db 617 PVLRVKAIIMEFNPREVARNVFCNDQVVKGEAGEVVRVCLHVQKSTRDLREGIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKSTRQTQVLTQTCETLKLQLPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKSTRQTQVLTQTCETLKLQLPNCIEDPVSPVILRLNF 736
QY 721 SLVGTPLSAFNGNLRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFNGNLRLPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSVYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVTVRNDGEDSVYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916
QY 901 QLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLFLVFPV 960
Db 917 QLELPVKYAVYVTVSHGVSTKYLNFTASENTSRVMQHQYQVSNLQGRSLPISLFLVFPV 976
QY 961 RLNOTVIWDRPOVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1020
Db 977 RLNOTVIWDRPOVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1036

1021 FPGIOEFNATLKGMLSPDWYIKTSHNHLLIYSTAILEFNDVSFTLLPGGAFVRSQTET 1080
1037 FPGIOEFNATLKGMLSPDWYIKTSHNHLLIYSTAILEFNDVSFTLLPGGAFVRSQTET 1096
1081 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
1097 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1153

RESULT 6
US-08-482-293A-3
Sequence 3, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482.293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
1 ENLDTENAMTFQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGCEPI 60
17 ENLDTENAMTFQENARGGQSVVQLQGSRRVVGAPQEIIVAAANQSGSLYQCDYSTGCEPI 76
61 RLQVPVAVNMSLGLSLAATTPPQLACGPTVHQTCSNTYVYKGLCFPGSNLRQOPQK 120
77 RLQVPVAVNMSLGLSLAATTPPQLACGPTVHQTCSNTYVYKGLCFPGSNLRQOPQK 136
121 FPEALRGCPQEDSIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 180

137 FPEALRGCPQEDSIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196
181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIKRELINITNGARKNAFKLILVI 240
197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKLILVI 256
241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNITASKPPRDHVFQVN 300
257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNITASKPPRDHVFQVN 316
301 NFEALKTIONQLREKIPALEGTQTSSSSPEHEMSQEGFSAAITSGNPILLSTVGSYDWDAG 360
317 NFEALKTIONQLREKIPALEGTQTSSSSPEHEMSQEGFSAAITSGNPILLSTVGSYDWDAG 376
361 GVFLYTSKEKSTFINNTRVDSDMNDAYLVYAAAIILNRNVQSLVGLGAPRYOHLGLVAMFR 420
377 GVFLYTSKEKSTFINNTRVDSDMNDAYLVYAAAIILNRNVQSLVGLGAPRYOHLGLVAMFR 436
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTREGGOVSCPL 480
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTREGGOVSCPL 496
481 PRGORARWQCDVLYGEGQPGWRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540
497 PRGORARWQCDVLYGEGQPGWRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 556
541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
601 PVLEVKAIMEFNPREVARNPVECDQVVKGEAGEVRVCLHVQKSTDRLEGOIQSVVT 660
617 PVLEVKAIMEFNPREVARNPVECDQVVKGEAGEVRVCLHVQKSTDRLEGOIQSVVT 676
661 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQETLKLQLPNCIEDPVSPIVRLNF 720
677 YDLALDSGRPHSRVAFNETKNSSTRQTVLGLTQETLKLQLPNCIEDPVSPIVRLNF 736
721 SLVGTPLSAFGNLPVLAEDAQRULFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFGNLPVLAEDAQRULFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796
781 GPREFNVTIVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRORSORSWRLACESASSTEV 840
797 GPREFNVTIVRNDGEDSYRTQVTFPPFDLSYRKVSTLQNRORSORSWRLACESASSTEV 856
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRINKTEF 900
857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRINKTEF 916
901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSLSPLVPLVPV 960
917 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHQYQVSNLQORSLSPLVPLVPV 976
961 RLNQTVIWDPRQVPTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSITAVCORIQCDIP 1020
977 RLNQTVIWDPRQVPTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSITAVCORIQCDIP 1036
1021 FPGIOEFNATLKGMLSPDWYIKTSHNHLLIYSTAILEFNDVSFTLLPGGAFVRSQTET 1080
1037 FPGIOEFNATLKGMLSPDWYIKTSHNHLLIYSTAILEFNDVSFTLLPGGAFVRSQTET 1096
1081 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
1097 KVEPFVFPNPLPLIVGSSVGGILLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1153

RESULT 7
US-08-943-363-3
Sequence 3, Application US/08943363
Patent No. 5837478
GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,363
FILING DATE: 5-AUG-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1153 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQBNARFGQGVVQLQGSRRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQBNARFGQGVVQLQGSRRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDAFLDGGSGIIPHPFRMKFEVSTVMQKKGKTLFSLIMQYSEEF 180
Db 137 FPEALRGCPQEDSDAFLDGGSGIIPHPFRMKFEVSTVMQKKGKTLFSLIMQYSEEF 196
QY 181 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILVI 240
Db 197 RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILVI 256
QY 241 TDGEKFGDPLGVEDVPEADRGVIRYIVGVGDAFPRSEKSRQELNTIASKPRDHVQFN 300
Db 257 TDGEKFGDPLGVEDVPEADRGVIRYIVGVGDAFPRSEKSRQELNTIASKPRDHVQFN 316
QY 301 NPEALKTIQNLREKIFAEGTQTGSSSFEHMSQEGFSAITNGPILLSTVGSDYDAG 360
Db 317 NPEALKTIQNLREKIFAEGTQTGSSSFEHMSQEGFSAITNGPILLSTVGSDYDAG 376

QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMPR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQSVCPCL 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQSVCPCL 496
QY 481 PRGORARWQCCDAVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWQCCDAVLYGEGQCPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTIVGAQGHVLLLSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTIVGAQGHVLLLSQ 616
QY 601 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY 661 YDLALDSGRPHSRAVFNETKNSTRROTQVGLGLTQTCETLKLQLENCIEDPVPVILRLNF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRROTQVGLGLTQTCETLKLQLENCIEDPVPVILRLNF 736
QY 721 SLVGTPLSAFNGRPLVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 780
Db 737 SLVGTPLSAFNGRPLVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 796
QY 781 GPREFNVTVRNDGEDSYETQVTFPPDLISYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYETQVTFPPDLISYRKVSTLQNRQSRQSWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSILPISLVLVPV 960
Db 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSILPISLVLVPV 976
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVVCNCSIAVCORIQCIP 1020
Db 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAEIRKAPVVCNCSIAVCORIQCIP 1036
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDSVFTLLPGGAFVRSQET 1080
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDSVFTLLPGGAFVRSQET 1096
QY 1081 KVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFFKRQKMDMSEGGPPGAEPO 1137
Db 1097 KVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFFKRQKMDMSEGGPPGAEPO 1153

RESULT 8
US-09-193-043-3
; Sequence 3, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6251395e1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match 99.9%; Score 5868; DB 3; Length 1153;									
Best Local Similarity 99.6%; Pred. No. 0;									
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFOENARFGQSVVQLQGRVVVVGAPQEIIVAA	NORGSLYQCDYSTGSC	60					
Db	17	FNLDTENAMTFOENARFGQSVVQLQGRVVVVGAPQEIIVAA	NORGSLYQCDYSTGSC	76					
Qy	61	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCS	ENTYVKGCLFGLGSLNRQ	120					
Db	77	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCS	ENTYVKGCLFGLGSLNRQ	136					
Qy	121	FPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKEFV	STVMEQKKSKTLFSLMOYSE	180					
Db	137	FPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKEFV	STVMEQKKSKTLFSLMOYSE	196					
Qy	181	RIHFTFKFQNNPNSLVKPIITQLGRTHATGVRKV	IRELLNITNGARKNAFKIL	240					
Db	197	RIHFTFKFQNNPNSLVKPIITQLGRTHATGVRKV	IRELLNITNGARKNAFKIL	256					
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVD	AFRSEKSRQELNTIASKPP	300					
Db	257	TDGKFGDPLGYEDVPEADREGVIRYVIGVD	AFRSEKSRQELNTIASKPP	316					
Qy	301	NFEALKTIONOLREKIFAEGTQGTSSSFHEHMS	QEGFSAATISNGPLLS	360					
Db	317	NFEALKTIONOLREKIFAEGTQGTSSSFHEHMS	QEGFSAATISNGPLLS	376					
Qy	361	GVFLYTSKEKSTFINMTVDSMDAYLGYAAAIL	LRNVOSLVLGAPYOHIGL	420					
Db	377	GVFLYTSKEKSTFINMTVDSMDAYLGYAAAIL	LRNVOSLVLGAPYOHIGL	436					
Qy	421	QNTGMWESNANVKGTOIGAYFGASLCSDVD	NSGSTDVLIGAPHYYE	480					
Db	437	QNTGMWESNANVKGTOIGAYFGASLCSDVD	NSGSTDVLIGAPHYYE	496					
Qy	481	PRGQARWQCDVAVLYGEOQPMGRFGAALT	VLGVNGDKLTDVAIG	540					
Db	497	PRGQARWQCDVAVLYGEOQPMGRFGAALT	VLGVNGDKLTDVAIG	556					
Qy	541	HGTSGGTSPISSHQSRIAGSKLSPRLQYFG	QSLSGGDLTMDGLV	600					
Db	557	HGTSGGTSPISSHQSRIAGSKLSPRLQYFG	QSLSGGDLTMDGLV	616					
Qy	601	PVLVRKALMEFNPREVARNVFCNDQVVK	GKEAGEVRVCLHVQK	660					
Db	617	PVLVRKALMEFNPREVARNVFCNDQVVK	GKEAGEVRVCLHVQK	676					
Qy	661	YDLALDSGPHGRVAFNETKNSTRQTVLGT	QTCETLKLQPN	720					
Db	677	YDLALDSGPHGRVAFNETKNSTRQTVLGT	QTCETLKLQPN	736					
Qy	721	SLVGTPLSAFGNLRPVLAEDAQLFTALP	FEKNCNDNICODDLS	780					
Db	737	SLVGTPLSAFGNLRPVLAEDAQLFTALP	FEKNCNDNICODDLS	796					
Qy	781	GPREENVTVVRNMGDSYRTQVTFPP	LDLSYRKVSTLQNR	840					
Db	797	GPREENVTVVRNMGDSYRTQVTFPP	LDLSYRKVSTLQNR	856					
Qy	841	SGALKSTCSINHPIIPENSEVTFNITFD	VDSKASLGNKLLK	900					
Db	857	SGALKSTCSINHPIIPENSEVTFNITFD	VDSKASLGNKLLK	916					
Qy	901	QLELPVKYAVYVMTVSHGVSTKYLNT	FTASNTSRVMQHQYQV	960					

Db	917	QLELPVKYAVYVMTVSHGVSTKYLNT	FTASNTSRVMQHQYQV	976					
Qy	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLP	SHSDFLAELRKAPV	1020					
Db	977	RLNQTVIMDRPQVTFSENLSSTCHTKERLP	SHSDFLAELRKAPV	1036					
Qy	1021	FFGQIEFNAFLKGNLSFDWYIKTSHNHL	LIVSTAEILFND	1080					
Db	1037	FFGQIEFNAFLKGNLSFDWYIKTSHNHL	LIVSTAEILFND	1096					
Qy	1081	KVPEPPEVNPPLPIVSGSSVGGLLLLAL	ITAAALYKLGFFK	1137					
Db	1097	KVPEPPEVNPPLPIVSGSSVGGLLLLAL	ITAAALYKLGFFK	1153					

RESULT 9

US-09-688-307A-3

Sequence 3, Application US/09688307A

Patent No. 6432404

GENERAL INFORMATION:

APPLICANT: Gallatin, Michael W.

APPLICANT: Van der Vieren, Monica

TITLE OF INVENTION: No. 6432404el Human Beta-2

FILE REFERENCE: 27866/36646

CURRENT APPLICATION NUMBER: US/09/688,307A

CURRENT FILING DATE: 2000-10-13

PRIOR APPLICATION NUMBER: 09/193,043

PRIOR FILING DATE: 1998-11-16

PRIOR APPLICATION NUMBER: 08/605,672

PRIOR FILING DATE: 1996-02-22

PRIOR APPLICATION NUMBER: 08/173,497

PRIOR FILING DATE: 1993-12-23

PRIOR APPLICATION NUMBER: 08/286,889

PRIOR FILING DATE: 1994-08-05

PRIOR APPLICATION NUMBER: 08/362,652

PRIOR FILING DATE: 1994-12-21

PRIOR APPLICATION NUMBER: 08/943,363

PRIOR FILING DATE: 1997-10-03

NUMBER OF SEQ ID NOS: 114

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 1153

TYPE: PRT

ORGANISM: Homo sapiens

US-09-688-307A-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFOENARFGQSVVQLQGRVVVVGAPQEIIVAA	NORGSLYQCDYSTGSC	60					
Db	17	FNLDTENAMTFOENARFGQSVVQLQGRVVVVGAPQEIIVAA	NORGSLYQCDYSTGSC	76					
Qy	61	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCS	ENTYVKGCLFGLGSLNRQ	120					
Db	77	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHQTCS	ENTYVKGCLFGLGSLNRQ	136					
Qy	121	FPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKEFV	STVMEQKKSKTLFSLMOYSE	180					
Db	137	FPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKEFV	STVMEQKKSKTLFSLMOYSE	196					
Qy	181	RIHFTFKFQNNPNSLVKPIITQLGRTHATGVRKV	IRELLNITNGARKNAFKIL	240					
Db	197	RIHFTFKFQNNPNSLVKPIITQLGRTHATGVRKV	IRELLNITNGARKNAFKIL	256					
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVD	AFRSEKSRQELNTIASKPP	300					
Db	257	TDGKFGDPLGYEDVPEADREGVIRYVIGVD	AFRSEKSRQELNTIASKPP	316					
Qy	301	NFEALKTIONOLREKIFAEGTQGTSSSFHEHMS	QEGFSAATISNGPLLS	360					

Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVOSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVOSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVQTRGGQVSVCP 496
Qy 481 PRGORARWOCDAVLVYGEQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLVYGEQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVLGTOTCETLKLQPNCTIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVLGTOTCETLKLQPNCTIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFSFMSLDCVLWG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFSFMSLDCVLWG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQVQVSNLQORSIPSLVFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMOHQVQVSNLQORSIPSLVFLVPV 976
Qy 961 RINQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1020
Db 977 RINQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCSTAVCQRIQCDIP 1036
Qy 1021 FFGIQEEFNATLKNLSFDWYIKTSHNLLIVSTABILFNDVSFTLLPGQAFVRSQDET 1080
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSHNLLIVSTABILFNDVSFTLLPGQAFVRSQDET 1096
Qy 1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFEKRYKQKMMSEGGPPGAEPQ 1137
Db 1097 KVEPFEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFEKRYKQKMMSEGGPPGAEPQ 1153

RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQSGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLQSGSRVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCCSENTYVVKGLCFLFGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCCSENTYVVKGLCFLFGSNLRQOPQK 136
Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRRMKFVSTVMEQKKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRRMKFVSTVMEQKKSKTLFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNPRLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILVI 240
Db 197 RIHFTFKFQNNPNPRLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILVI 256
Qy 241 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAPFSEKSRQBELNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAPFSEKSRQBELNTIASKPPRDHVFQVN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITSNGLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVOSLVGLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVOSLVGLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVQTRGGQVSVCP 496
Qy 481 PRGORARWOCDAVLVYGEQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLVYGEQGPWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKSTRQTVLGTOTCETLKLQPNCTIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVLGTOTCETLKLQPNCTIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFSFMSLDCVLWG 780
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFSFMSLDCVLWG 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 916
901 QLELPVKYAVYVMTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 960
917 QLELPVKYAVYVMTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 976
961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PNVNCSIAVCORIQCDIP 1020
977 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PNVNCSIAVCORIQCDIP 1036
1021 FFGIQEEFNATLKGNSFDWYIKTSHNHLIVSTAELLFNDVSFTLLPGQGA FVRSQTET 1080
1037 FFGIQEEFNATLKGNSFDWYIKTSHNHLIVSTAELLFNDVSFTLLPGQGA FVRSQTET 1096
1081 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137
1097 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153

RESULT 11

US-08-476-062A-43
Sequence 43, Application US/08476062A
Patent No. 5877275
GENERAL INFORMATION:
APPLICANT: Ainaout, M. Amin
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,062A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/216,081
FILING DATE: 21-MAR-1994
APPLICATION NUMBER: 07/637,830
FILING DATE: 04-JAN-1991
APPLICATION NUMBER: 07/539,842
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: 07/212,573
FILING DATE: 28-JUN-1988
ATTORNEY/AGENT INFORMATION:
NAME: Freeman, John W.
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/068003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1152 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-476-062A-43

Query Match 99.4%; Score 5837.5; DB 2; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

1 FNLDTENAMTFQENARGFGQSVVQGGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 60
17 FNLDTENAMTFQENARGFGQSVVQGGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCPEI 76
61 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120
77 RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136
121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEFVSTVMEQLKSKTLLFSLMQYSEEF 180
137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEFVSTVMEQLKSKTLLFSLMQYSEEF 196
181 RIHFTFKFQNNPNPSLVKPIQTLLGRTHATATGVRKVIKRELLININGAKNAFKLILVI 240
197 RIHFTFKFQNNPNPSLVKPIQTLLGRTHATATGVRKVIKRELLININGAKNAFKLILVI 256
241 TDGEKFGDPLGYEDVITPEADREGVIRVYGVGDAFSEKSRQELINTIASKPPDRHVFQVN 300
257 TDGEKFGDPLGYEDVITPEADREGVIRVYGVGDAFSEKSRQELINTIASKPPDRHVFQVN 316
301 NFEALKTIQNOLREKIFAIEGTOTGSSSSPEHEMSQEGFSAAITSGNPLLTSTVGSYDWAG 360
317 NFEALKTIQNOLREKIFAIEGTOTGSSSSPEHEMSQEGFSAAITSGNPLLTSTVGSYDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMER 420
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMER 436
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVVCP 480
437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVVCP 496
481 PRGORARWQCDVLYGQGGPQWGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAVYLF 540
497 PRG-RARWQCDVLYGQGGPQWGRFGAALTVLGVDVNGDKLTDVAIGAPEGEDNRGAVYLF 555
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVILLRSQ 600
556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQHVILLRSQ 615
601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660
616 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 675
661 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQLPNCIEDPVPVILRLNF 720
676 YDLALDSGRPHSRVAFNETKSTRRTQVGLGTCTETLKLQLPNCIEDPVPVILRLNF 735
721 SLVGTPLSAFONLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITTFSEMSLCLVVG 780
736 SLVGTPLSAFONLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITTFSEMSLCLVVG 795
781 GPREFNVTVVRNDGEDSYRTQVTFEFPPLDLSYRKVSTLQNRQORSWRLACESASSTEV 840
796 GPRESNVTVVRNDGEDSYRTQVTFEFPPLDLSYRKVSTLQNRQORSWRLACESASSTEV 855
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900
856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 915
901 QLELPVKYAVYVMTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 960
916 QLELPVKYAVYVMTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 975
961 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PNVNCSIAVCORIQCDIP 1020
976 RLNQTVWDRPQVTFSENLSSTCHTKERLPSSHDFLAELRKA PNVNCSIAVCORIQCDIP 1035
1021 FFGIQEEFNATLKGNSFDWYIKTSHNHLIVSTAELLFNDVSFTLLPGQGA FVRSQTET 1080
1036 FFGIQEEFNATLKGNSFDWYIKTSHNHLIVSTAELLFNDVSFTLLPGQGA FVRSQTET 1095
1081 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137

```
Db 1096 KVEPFEVNPDLPIVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1152
|||||
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaout
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 99.4%; Score 5837.5; DB 5; Length 1152;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGLYQCDYSTGSCBPI 60
|||
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGLYQCDYSTGSCBPI 76
|||
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPOK 120
|||
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVVKGLCFLFGSNLRQQPOK 136
|||
QY 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKBFVSTVMQKSKTLFSLMQYSEEF 180
|||
Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKBFVSTVMQKSKTLFSLMQYSEEF 196
|||
QY 181 RIHFTFKFQNNPNSRLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 240
|||
Db 197 RIHFTFKFQNNPNSRLVKPITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILVI 256
|||
QY 241 TDGEXFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQVN 300
|||
Db 257 TDGEXFGDPLGYEDVPEADREGVIRYVIGVDGDAFRSEKSRQELNTIASKPPRDHVFQVN 316
|||
QY 301 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSNGPLLSVGSYDWAG 360
|||
Db 317 NFEALKTIQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITSNGPLLSVGSYDWAG 376
|||
```

```
QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALILRNVRQSVLVLGAPRYOHIGLVAMFR 420
|||
Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALILRNVRQSVLVLGAPRYOHIGLVAMFR 436
|||
QY 421 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 480
|||
Db 437 QNTGMWESNANVKGTGIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 496
|||
QY 481 PRQEARWCQDAVLYGEGQGPWGRFCAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540
|||
Db 497 PRG-RARWCQDAVLYGEGQGPWGRFCAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 555
|||
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGODLTMDGLVDLTGVAQGHVLLRSQ 600
|||
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGODLTMDGLVDLTGVAQGHVLLRSQ 615
|||
QY 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLREGIOISVVT 660
|||
Db 616 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLREGIOISVVT 675
|||
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQCTETLKLQLPNCIEDPVSPIVLRNF 720
|||
Db 676 YDLALDSGRPHSRVAFNETKSTRQTQVLGTLTQCTETLKLQLPNCIEDPVSPIVLRNF 735
|||
QY 721 SLVGTPLSAFGNLRPVLAEDAORLFTALPFFPKNCNDNICQDDLSITFSFMSLDCLVVG 780
|||
Db 736 SLVGTPLSAFGNLRPVLAEDAORLFTALPFFPKNCNDNICQDDLSITFSFMSLDCLVVG 795
|||
QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPFDLDSYRKVUSTLONORSQBSWELACESASSTEV 840
|||
Db 796 GPREFNVTVTVRNDGEDSVRTQVTFPFDLDSYRKVUSTLONORSQBSWELACESASSTEV 855
|||
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPRTNKTEF 900
|||
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANVTSENNMPRTNKTEF 915
|||
QY 901 QLELPVKYAVYVWVTSHGVSSTKYLNFNTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 960
|||
Db 916 QLELPVKYAVYVWVTSHGVSSTKYLNFNTASENTSRVMQHOYQVSNLQGRSLPISLVFLVPV 975
|||
QY 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSITAVCQRCQDIP 1020
|||
Db 976 RLNQTVIWDROPVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNVNCSITAVCQRCQDIP 1035
|||
QY 1021 PFGIQEEFNATLKGNSLFDWIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQET 1080
|||
Db 1036 PFGIQEEFNATLKGNSLFDWIKTSHNHLIIVSTAEILFNDVSFTLLPQCGAFVRSQET 1095
|||
QY 1081 KVEPFEVNPDLPIVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137
|||
Db 1096 KVEPFEVNPDLPIVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1152
|||
```

RESULT 13

```
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2
; LENGTH: 1152
5424399-2
```

Query Match 99.4%; Score 5837.5; DB 6; Length 1152;

Best Local Similarity 99.4%; Pred. No. 0;									
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;									
Qy	1	FNLDTENAMTFOENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQORGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQORGSLYQCDYSTGSCPEI	76						
Qy	61	RLOVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK	120						
Db	77	RLOVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK	136						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	180						
Db	137	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	196						
Qy	181	RIHFTFKFQNNPNRSLVKPIITQLLGRTHATGVRKVIRELININGARKNAFKILIVI	240						
Db	197	RIHFTFKFQNNPNRSLVKPIITQLLGRTHATGVRKVIRELININGARKNAFKILIVI	256						
Qy	241	TDGEKGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTIASPPRDHVPQVN	300						
Db	257	TDGEKGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTIASPPRDHVPQVN	316						
Qy	301	NFALKTIQNLREKIPAEIGTGTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG	360						
Db	317	NFALKTIQNLREKIPAEIGTGTGSSSFEHMSQEGFSAATISNGPLLSTVGSYDWAG	376						
Qy	361	GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSVLGPAPRYOHIGLVAMFR	420						
Db	377	GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSVLGPAPRYOHIGLVAMFR	436						
Qy	421	QNTGMESNANVKTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSVCP	480						
Db	437	QNTGMESNANVKTQIGAVFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSVCP	496						
Qy	481	PRGORARWQCDVLYIGQGQPGWGFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYL	540						
Db	497	PRG-RARWQCDVLYIGQGQPGWGFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYL	555						
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	600						
Db	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	615						
Qy	601	PVLRYKAIMENPREVARNFECDQVVKGEAGEVVCVCHVOKSTDRREGQIQSVVT	660						
Db	616	PVLRYKAIMENPREVARNFECDQVVKGEAGEVVCVCHVOKSTDRREGQIQSVVT	675						
Qy	661	YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	720						
Db	676	YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	735						
Qy	721	SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCMDNICQDDLSITFSFMSLDCLVWG	780						
Db	736	SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCMDNICQDDLSITFSFMSLDCLVWG	795						
Qy	781	GPREFNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNRORSORSLACESASSTEV	840						
Db	796	GPRESNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNRORSORSLACESASSTEV	855						
Qy	841	SGALKSTSCSINHIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT	900						
Db	856	SGALKSTSCSINHIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKT	915						
Qy	901	QLELPVKYAVMVTSGVSTKYLNFTASENSTRVMQHOYQVSNLQORSLSISVFLVPV	960						
Db	916	QLELPVKYAVMVTSGVSTKYLNFTASENSTRVMQHOYQVSNLQORSLSISVFLVPV	975						
Qy	961	RLNQTVIWRQVTFSENLSTCHTKERLPSHDFLAELRKAPVVCNSIAVCQRIQCDIP	1020						
Db	976	RLNQTVIWRQVTFSENLSTCHTKERLPSHDFLAELRKAPVVCNSIAVCQRIQCDIP	1035						
Qy	1021	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRQ	1080						

Db	1036	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRQ	1095						
Qy	1081	KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ	1137						
Db	1096	KVEPPEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ	1152						
RESULT 14									
US-08-476-062A-44									
; Sequence 44, Application US/08476062A									
; Patent No. 5877275									
; GENERAL INFORMATION:									
; APPLICANT: Arnaout, M. Amin									
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY									
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS									
; NUMBER OF SEQUENCES: 53									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: Fish & Richardson P.C.									
; STREET: 225 Franklin Street									
; CITY: Boston									
; STATE: MA									
; COUNTRY: US									
; ZIP: 02110-2804									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Diskette									
; COMPUTER: IBM Compatible									
; OPERATING SYSTEM: Windows95									
; SOFTWARE: FastSeq for Windows Version 2.0									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/08/476,062A									
; FILING DATE: 07-JUN-1995									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER: 08/216,081									
; FILING DATE: 21-MAR-1994									
; APPLICATION NUMBER: 07/637,830									
; FILING DATE: 04-JAN-1991									
; APPLICATION NUMBER: 07/539,842									
; FILING DATE: 18-JUN-1990									
; APPLICATION NUMBER: 07/212,573									
; FILING DATE: 28-JUN-1988									
; ATTORNEY/AGENT INFORMATION:									
; NAME: Freeman, John W.									
; REGISTRATION NUMBER: 29,066									
; REFERENCE/DOCKET NUMBER: 00786/068003									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: 617/542-5070									
; TELEFAX: 617/542-8906									
; TELEX: 200154									
; INFORMATION FOR SEQ ID NO: 44:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 1163 amino acids									
; TYPE: amino acid									
; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
US-08-476-062A-44									
Query Match 59.1%; Score 3473; DB 2; Length 1163;									
Best Local Similarity 61.3%; Pred. No. 4, 2e-285;									
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;									
Qy	1	FNLDTENAMTFOENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQORGSLYQCDYSTGSCPEI	60						
Db	20	FNLDTBELTAFRVDSAGFGDSVQYANSVVVVGAPQKITAANQTGGLYQCYSYTGACEPI	79						
Qy	61	RLOVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK	120						
Db	80	GLQVPPPAVNMVSLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCLLQPT--QLTQR	137						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	180						
Db	138	LPVSRQCPQEGQDIVFLIDGSGISSENFAFMFVRAVISQFQRPSTQSLQFSNKF	197						
Qy	181	RIHFTFKFQNNPNRSLVKPIITQLLGRTHATGVRKVIRELININGARKNAFKILIVI	240						

